Appendix G

Geo-Neutral Point Source Model

General Comments

1. Appendix G and Geo-Neutral Point Source Model Evaluation Overview notes; Assessment of RI Geo-neutral Newtown Creek and Bowery Bay InfoWorks models: The InfoWorks model used in the RI to generate flows from the NYC collection system to Newtown Creek is generally consistent with the NYC model, and reasonably produces the results presented in their report and in the digital files provided to EPA. As the exact model was not provided, and the version of InfoWorks (ICM 7.5) used by EPA to check the model likely differed from the older InfoWorks CS version used for the RI, small differences observed in the results were expected. In addition to the numbered comments presented below, identified issues include:

    a. Hourly data were used for the simulations. The models were calibrated to 5-minute data. Use the same data frequency as input for simulation of CSO discharges to receiving waters.

    b. The Bowery Bay model EPA reviewed did not run without a few minor modifications. As the model appears to have been prepared by NYC and merely passed along via the NCG, this issue was likely due to an integrity check in the newer software that was not present in older versions of the software

2. Appendix G and Geo-Neutral Point Source Model Evaluation, Precipitation Data for Newtown Creek InfoWorks model: EPA will convene a meeting with technical representatives of the NCG, NYCDEP, and NYSDEC (similar to the modeling working group meetings) within the next 30 days to identify the appropriate precipitation data set to be used in the Geo-Neutral Point Source Model for the Newtown Creek site as part of the CERCLA RI/FS process. EPA will provide additional communications regarding this meeting within 5 business days of this comment transmittal.

3. Appendix G and Geo-Neutral Point Source Model Evaluation, discrepancy between meteorological stations and comparable daily datasets: For some years, there is a discrepancy between the hourly data reported for the meteorological stations and the comparable daily datasets. The daily data are generally the most reliable. The hourly data were frequently deficient in the early years of the ASOS program from the late 1990s through the early 2000s. Hourly data at LaGuardia from 1996-2005 in some years is 6 percent less than the reported daily totals. This discrepancy has largely been rectified more recently, but the 2012 hourly dataset is also deficient, with 3 inches less precipitation than its daily counterpart. Base CSO simulations with the objective of representing loads to the receiving waters on precipitation datasets adjusted to incorporate the reported daily precipitation totals.

Geo-Neutral Point Source Table 1 (included at the end of this section) presents annual precipitation for Central Park, LaGuardia, and JFK from daily datasets, along with annual precipitation at LaGuardia from its hourly dataset.

4. Appendix G and Geo-Neutral Point Source Model Evaluation, precipitation data frequency: As the RI's goal was to produce best estimates of CSO discharges, use the same precipitation timestep as was used in model calibration. Using data with a longer timestep results in smaller peak runoff rates, and thus underestimates CSO. Various means are available for obtaining sub-hourly data. Since the early 2000s, each principal weather station has recorded 1-minute data via the Automated Surface Observing System (ASOS). These data are not quality-controlled by NOAA, but are available to the public, and are generally of good quality. Hourly data can also be synthetically disaggregated to develop high frequency datasets that reproduce the variability expected in short-duration measurements, or other stations in the area could have been used to inform development of long-term 5-minute time series.

5. Appendix G and Geo-Neutral Point Source Model Evaluation, evaporation: Appendix G of the RI states: "*Daily evapotranspiration data were obtained from the Northeast Regional Climate Center at Cornell University. The NCB portion of the 2015 geo-neutral point source model used LGA evapotranspiration data for 1999 and CPK data for the 2000 to 2012 period because CPK evapotranspiration data were not available before 2000. The BBL portion of the 2015 geoneutral point source model used LGA evapotranspiration data for the entire 14-year period*." However, no evapotranspiration data for either site are reported to the National Weather Service. While pan evaporation was measured at Central Park from 1944-1958, evapotranspiration is usually derived from measurements of air temperature, solar radiation, vapor pressure, and wind speed. InfoWorks requires free surface evaporation as input; this value can be considered the same as potential evapotranspiration (PET) for this modeling. PET is usually greater than actual evapotranspiration.

   The InfoWorks input data indicate annual average evaporation of 28.9 inches (736 mm) for the Newtown Creek area, and 26.2 inches for Bowery Bay. While the methodology is not discussed in the report, it is likely NRCC's adaptation of a PET model for a grass-covered surface described on the Cornell website (www.nrcc.cornell.edu/wxstation/pet/pet.html), as the results nearly match monthly averages presented there for LaGuardia. Other methods of estimating PET or free surface evaporation yield higher annual averages:

   a. NWS atlases 33 and 34 (Farnsworth et al., 1982) present pan evaporation estimates nationwide, and coefficients for converting these estimates to free surface evaporation. Annual pan estimates for LaGuardia and Newark are 54.55 and 49.69 inches, and the conversion coefficient for the area is 0.78, yielding free surface evaporation of 42.5 in/y and 38.8 in/y, respectively.

   b. Vogel and Sankarasubramanian (2015) present PET estimates for 1,469 sites nationwide based on the Hargreaves-Samani method. The nearest sites in their dataset are Chatham NJ (20 miles west of Newtown Creek) and Mahwah NJ (30 mi NW) with respective estimates of 41.4 inches/year and 40.0 in/y.

c. Application of the Hargreaves-Samani (1985) method using 1999-2012 Central Park and LaGuardia daily temperatures for Newtown Creek (latitude 40.74°N) yields 36.2 in/y and 35.8 in/y, respectively.

d. Application of the Hamon (1961) method for free surface evaporation based on Central Park and LaGuardia daily temperatures for 1999-2012 at Newtown Creek (latitude 40.74°N) yields 30.7 in/y and 31.8 in/y, respectively.

The estimates used in InfoWorks thus appear low. Additionally, the rationale for the 10 percent difference between Newtown Creek and Bowery Bay is not apparent. While evaporation is a small component of the water balance in urban runoff, and the impact of its underestimation is likely small, underestimation of evaporation yields slightly more runoff, and thus likely slightly overestimates CSO. Apply the Hargreaves-Samani method with LaGuardia daily temperatures for determining daily evaporation boundary conditions.

Table 1. Annual Precipitation (inches)

Year	JFK Daily	NYC Daily	LGA Daily	LGA Hourly
1990	45.24	60.92	51.22	51.32
1991	38.73	45.18	38.16	38.16
1992	38.38	43.35	37.40	37.40
1993	35.61	44.28	43.16	43.16
1994	43.33	47.39	43.49	43.45
1995	34.42	40.42	35.31	35.35
1996	51.45	56.19	49.12	46.12
1997	39.87	43.93	45.37	45.30
1998	37.55	48.69	45.21	44.32
1999	40.10	41.51	41.07	39.80
2000	41.02	45.45	42.48	39.84
2001	32.72	35.65	33.97	32.07
2002	43.13	45.20	44.84	42.01
2003	44.77	58.42	54.96	51.82
2004	50.95	51.93	50.68	49.61
2005	49.55	55.97	48.16	45.40
2006	44.80	59.89	53.95	53.95
2007	46.91	61.67	53.43	53.43
2008	46.26	53.61	47.84	47.79
2009	45.88	53.62	46.33	46.26
2010	42.47	49.37	40.63	40.30
2011	55.78	72.81	65.34	65.33
2012	39.85	38.51	36.71	33.23
2013	35.48	46.32	38.29	38.14
2014	50.75	53.79	50.31	50.08
2015	38.31	40.97	37.20	38.55
2016	36.01	42.17	39.39	37.89
Average	42.57	49.53	44.96	44.08

Table 1 Data Sources

1. NOAA 2016 Local Climatological Data Annual Summary with Comparative Data New York, JFK International Airport (KJFK), National Centers for Environmental Information, Asheville, NC www.ncdc.noaa.gov/IPS/lcd/lcd.html

2. NOAA 2016 Local Climatological Data Annual Summary with Comparative Data New York, La Guardia Airport (KLGA), National Centers for Environmental Information, Asheville, NC www.ncdc.noaa.gov/IPS/lcd/lcd.html

3. NOAA 2016 Local Climatological Data Annual Summary with Comparative Data New York, New York, New York (KNYC), National Centers for Environmental Information, Asheville, NC www.ncdc.noaa.gov/IPS/lcd/lcd.html

4. NOAA, 2016. LGA hourly 1990-2013: www.ncdc.noaa.gov/cdo-web/search?datasetid=PRECIP_HLY#

5. NOAA, 2016. LGA hourly 2014-2016: www.ncdc.noaa.gov/qclcd/QCLCD

Specific Comments

1. Figure G3-19. The value for 2012 appears to be incorrect; it should be 38.5 inches, not 49.5 inches. Verify and correct the value.

2. Page 31, Section 3.5.1 Diagnostic Analysis of 2015 Geo-Neutral Point Source Model. Present annual precipitation 2008-2012 for Central Park (NYC) and LaGuardia Airport (LGA) in a table. The values from the Global Historical Climate Network (GHCN) daily dataset are provided below.

Year	NYC	LGA
2008	53.61	47.84
2009	53.62	46.33
2010	49.37	40.63
2011	72.81	65.34
2012	38.51	36.71
Average	53.58	47.37
Min	38.51	36.71
Max	72.81	65.34

3. Page 35, Section 3.6, Model Application:

   a) LGA precipitation averages given as 46.0 for 1999-2012 and 44.4 for 1980-2012. Review of the rainfall records indicate 47.2 and 44.9 for these same periods, respectively, from the GHCN daily dataset. Verify that the values are correct or explain the discrepancy.

   b) Explain the significance of using 1980-2012, and the 14- and 50-year periods used for Central Park. Uninterrupted daily records for NYC begin in 1876, and for LGA in 1944.

4. Page 36, Section 3.6 Model Application, third paragraph. Evapotranspiration data for either the Northeast Regional Climate Center at Cornell or the La Guardia location could not be located for recent decades. Cornell likely provided estimates based on a method such as Penman-Monteith. It is not clear why such estimates would be available for only part of the time period. Clarify the source and type of evapotranspiration data should be clarified.

References

Farnsworth, R.K., Thompson, E.S., and Peck, E.L, 1982, Evaporation atlas for the contiguous 48 United States, NOAA Technical Report NWS 33, National Oceanic and Atmospheric Administration, Washington, DC, p. 27.

Farnsworth, R.K. and Thompson, E.S., 1982, Mean monthly, seasonal, and annual pan evaporation for the United States, NOAA Technical Report NWS 34, National Oceanic and Atmospheric Administration, Washington, DC, p. 85.

Hamon, W.R., 1961, Estimating potential evapotranspiration: Journal of Hydraulics Division, Proceedings of the American Society of Civil Engineers, v. 87, p. 107–120.

Hargreaves, G. H., and Samani, Z. A. (1985). "Reference crop evapotranspiration from temperature." Appl. Eng. Agric., 1(2), 96–99.

Vogel, R.M., and A. Sankarasubramanian. 2015. Monthly Climate Data for Selected USGS HCDN Sites, 1951-1990, R1. ORNL DAAC, Oak Ridge, Tennessee, USA.

Hydrodynamic and Sediment Transport Models

General Comments

1. Review of the graphics and text describing the tide boundary condition at the northern boundary suggests that the model input was the result of a calibration exercise. The use of a calibrated boundary condition is not standard practice and is not technically defensible. In addition, using the calibrated boundary instead of the correct water levels has an impact on water levels and currents in the project area. Given these arguments, EPA strongly recommends the use the Lower Passaic River and Newark Bay Superfund or NYC LTCP regional model (or outputs of one of these same models) to specify the tide at the northern boundary; boundary conditions for temperature and salinity can also be specified using the outputs of the selected regional model.

   Furthermore, not driving the hydrodynamic and salinity transport model with a regional model propagates unnecessary uncertainties into both the sediment and contaminant transport models. If this change is not made, EPA strongly recommends that an independent assessment be made to quantify the impacts of using the extrapolated boundary conditions in the East River on the transport of both water and salinity, as well as using these results to simulate the transport of sediments and contaminants.

2. Although a tremendous amount of work has gone into the development of the sediment transport model, the values of certain parameters (e.g., settling velocity of the fine sediment size class) required the use of values that are not usually measured for flocculated sediments in estuarine waters. EPA recommends that additional validation is performed (e.g., showing comparisons between simulated and measured suspended sediment concentration profiles at different locations in Newtown Creek under different tidal and runoff conditions). As is, and considering the issue mentioned in General Comment No. 3 below, EPA is concerned about the accuracy of the sediment transport model when used for driving the chemical fate and transport model. EPA strongly recommends that the model and report be revised accordingly.

3. Although a diagnostic analysis was performed for the simplified representation of sediment transport in the East River, it does not appear that the impact on sediment transport in the East River has been thoroughly assessed. This concern will increase with the use of chemical fate and transport model to simulate the transport and fate of sorbed contaminants on sediments that are being transported out of Newtown Creek into the East River and vice versa. EPA is concerned that the use of a simplified sediment transport model for the East River to represent the transport and fate of sediments and sorbed contaminants exchanged between the East River and Newtown Creek will not accurately represent these processes.

4. Propwash Resuspension Submodel: Revisions to propwash resuspension submodel described in Section 2.2.2.4 of the Final Modeling Results Memorandum (FMRM) represent a vast improvement to the first version of the submodel.

5. Verification of Model Inputs: Verification of the model inputs could not be performed because the input files for this submodel were not provided. The goal of this task is to insure the inputs were correctly specified in the input files. Provide these inputs when they are ready.

6. Verification of Model Calculations: The calculations of the propwash resuspension submodel were checked by reviewing the model code to verify that the submodel computes bed scour due to propwash correctly as given in Section of Attachment G-K (Details of Propwash Resuspension Submodel Structure and Formulation). The finding from this task was that the code in the sed_sedflx_SEDZLJ.f subroutine correctly represented the equations for propeller thrust, velocities in the zone of flow establishment and the zone of established flow, and the calculation of bed scour due to propwash. The novel subgrid approach used to calculate scour within a grid cell due to propwash from a moving ship is impressive.

   However, as stated above, the lack of input files did not allow verification of the parameters and variables used in the calculation of bed scour due to propwash. As a result, it was not possible to verify that correct values for the parameters are being used in the calculations and that variables in this submodel are being calculated correctly.

7. Benchmarking of Model Outputs: The lack of input files did not allow verification of integrity of output from the model by recompiling the source code, re-running the one- year simulation in which the propwash submodel was activated with the generated code executable, and comparing the model results from this simulation to the results (as described by Hayter [2016]).

   Provide all input files for the propwash resuspension submodel prior to submittal of the chemical fate and transport model to EPA for review.

8. The sediment mass loss associated with the simulated tide-induced wetting and drying during a one-year model run was investigated to determine if the use of HDRY = 0.1 m and HWET = 0.13 m was satisfactorily mass conserving. The results of this evaluation determined that sediment mass loss was minimal, and thus the model did satisfactorily conserve sediment mass. This check is important in models in which simulated wetting and drying occurs since all wetting and drying routines are relatively crude approximations that are not based on first principles of mass, momentum and energy conservation.

9. The FMRM should be a comprehensive documentation of the modeling study. Currently, it is structured partially as a document describing the refinements to the PMRM model (for example, see the discussion in Sections 4.1, 5.1 and 7.2). Either (1) include the PMRM as an attachment to the FMRM, or (2) simplify the FMRM text to discuss only the final model framework/formulations.

10. The continuous salinity data used for model calibration are not synoptic with the rest of the hydrodynamic data and are also available only for three months during a relatively dry period. The report indicates that nine months of synoptic continuous data were determined to be unreliable due to sonde calibration issues. Present a more detailed explanation of this problem (and how it can be avoided in the future) and describe how it affected sondes at all the locations

for the complete 9-month period (e.g., this issue limits the validation of the model during large point source discharge events).

11. Hydrodynamic model calibration is discussed in the text using one single statistical value per variable. This results in grouping of all the information from different stations, environmental conditions (dry- and wet-weather), etc. Present a more detailed evaluation of the model performance with statistical evaluation for individual locations and during specific conditions that are important for the project. The statistical evaluation should be performed and discussed in the text for the individual locations, using the metrics of bias and ubRMSD already included in the FMRM, as well as relative bias. Since the potential application of this model includes testing various remedial/management strategies, including source control, present an evaluation of model performance during specific environmental conditions such as dry-weather and large point source discharge events. In addition, model-data comparisons in tidal environments are typically performed by comparing model results to the measured amplitude and phase of various tidal constituents. Include such a quantitative comparison of the tidal constituents in the evaluation (this is applicable to both water level and currents).

12. The FMRM document is missing an analysis/discussion of the dominant fate and transport processes for sediments within Newtown Creek which need to be reproduced by the numerical model. Specific questions to be addressed include:

    a. What are the fate and transport processes evident in the data?

        i. Over tidal timescales during dry-weather conditions
        ii. During wet-weather conditions

    b. What processes are important for fate and transport and need to be represented in the model?

        i. How important is erosion and deposition of sediments under both wet weather and dry weather tidal conditions?
        ii. Does erosion not occur under normal tidal conditions, as the model currently suggests? Is this consistent with what is happening in the Creek?
        iii. How important is navigation scour for fate and transport of sediments? Is it locally important (e.g., formation of scour holes), or is it globally important?
        Such an analysis and discussion will ensure that relevant fate and transport processes have been appropriately incorporated into the model framework, and provide confidence in model projections for the future. Revise the document to include these analyses/discussions.

13. NSRs represent the only calibration metric in the FMRM sediment transport model application. As such, a number of datasets have been analyzed to support evaluation of the calibration – bathymetric differencing (1991-2012, 1999-2012, 1999-2011, and 1991-1999), geo-chronology cores, and historical dredging records. These analyses are presented in two separate Attachments (G-G and G-H). The results of these analyses are presented in Figure G5-5 and G5-6 without attempting to reconcile what seems, at first glance, very different NSRs between methodologies. For instance, the various lines of evidence for NSRs in Maspeth Creek vary by approximately one order of magnitude (~0.75 cm/yr to 7 cm/yr). Revise the report to include

an analysis and associated discussion of the NSRs from the various lines of evidence. Per EPA's review, accounting for navigation history and sources of bias in the bathymetry data and geochronology cores, the various lines of evidence tend to roughly similar conclusions on the current NSRs in the various tributaries. That is an important conclusion that indicates consistency amongst the various lines of evidence, and strengthens the resulting NSR calibration metric for the numerical model. Revise the document to state this.

14. Some of the assumptions and statements in the FMRM document are either not presented, or presented without adequate evidence and justification. For instance, Section 5.2.2 Data-Based Mass Balance Analysis, includes an implicit assumption of no deposition in the tributaries of solids originating from the East River and the Main Stem, with no overt mention in the text. The same section also includes the statement "more than 90% of tributary deposition is due to point source sediment loads". However, no evidence or discussion is provided in support of this statement. Revise the text so that all assumptions and statements are explicitly listed, justified, and discussed.

15. Several figures are presented in the text without adequate explanation of the information in the graphics, nor a presentation/discussion of the conclusions from the graphics. Specific examples include Figures G5-5, G5-6, G5-28, G5-36, etc. Revise the text with adequate description and discussion of the information presented in each figure and associated conclusions.

16. The notion of temporal decline in point source loadings over time is currently presented in several places, e.g., Attachment G-G, Attachment G-I, Appendix G Section 5.2.1, etc. Given the potential importance of this topic to the historical evolution of the study area, revise the document to include a separate section or attachment exploring this hypothesis, and presenting the various lines of both direct and indirect evidence.

17. The TSS boundary condition at the East River boundary is defined in a relatively simplistic manner, as a temporal and vertical-average value. This approach neglects potential seasonality in TSS and vertical gradients that are relevant in the presence of estuarine circulation. Review the TSS data for temporal (seasonal as well as spring-neap) and vertical gradients and incorporate into the model boundary conditions, as appropriate. This will result in a TSS boundary condition that is better constrained and may help improve model-data comparisons for TSS.

18. The FMRM model application involves a relatively large number of calibration parameters compared to calibration metric (only NSR). Calibration parameters include:

    a. Wash load fraction of East River solids load

    b. Flocculated clay/silt fraction of East River solids load

    c. Fine sand fraction of East River solids load

    d. Settling velocity of flocculated clay/silt from East River

    e. Settling velocity of flocculated clay/silt from point sources

f. If including propeller scour, then

   i. Settling velocity of scoured cohesive sediments
   ii. Probability of resuspension

This is a relatively large number of calibration parameters (n=7) compared to the number of calibration metrics (n=1). The primary concern generated by this comparison is the possibility of obtaining non-unique solutions. In other words, there may be multiple combinations of parameter values that can result in a good model performance relative to the sole calibration metric. In addition, the settling velocity of the wash load from East River is based on an assumed value. EPA strongly recommends developing data-based methodologies to reduce the number of calibration parameters. This will lead to unique parameter values, and inputs that are data-based and technically defensible.

19. In general, several sediment transport model inputs and parameters that should be treated as model input are either assumed (e.g., particle diameters for the fine and medium-coarse sand classes, settling velocity of wash load) or are subject to calibration (the wash load, flocculated clay/silt, and fine sand fractions of suspended sediment entering at the East River boundaries). For instance, the particle diameters of the fine and medium-coarse sand fractions can be determined from bed grain size distribution measurements conducted as part of the RI program. Similarly, the various size fractions at the East River boundaries should be based on measurements as was done for the point source loadings. The settling velocity of the wash load fraction can be calculated using Stokes Law based on the measured particle diameters and specific gravities associated with this size class. EPA recommends revising the model to parameterize inputs using site-specific data as described above to minimize the potential for model artifacts that may arise from assumed/calibrated inputs.

20. The FMRM sediment transport model application has been calibrated to a single metric (NSRs). This approach can result in a biased model if the calibration metric also happens to be biased or affected by some artifact. The typical approach for sediment transport model applications for Superfund as well as other environmental applications is to calibrate to multiple lines-of-evidence. Such an approach will facilitate identification of biases in individual datasets (if such biases do not affect all metrics) and allow these biases to be suitably addressed as part of the model calibration. Additional calibration metrics for Newtown Creek include TSS measurements from water samples, TSS time-series estimated from the bulkhead turbidity measurements during Phase 2, limited TSS time-series estimated from Acoustic Backscatter (ABS) measurements by ADCPs during Phase 1, and limited suspended sediment fluxes using ABS data. Establishing model calibration over several metrics will allow calibration over various spatial and temporal scales and ensure that the resulting model performance is more robust and more rigorously tested. In addition, reviewing model results relative to TSS time-series data will also demonstrate model performance over varying time-scales and environmental conditions, e.g., tidal timescales (dry-weather), wet-weather conditions, navigation scour events, etc. EPA strongly recommends revising the model calibration strategy to use such a multiple lines-of-evidence approach to establish model calibration.

21. The FMRM model includes the application of detailed mechanistic sub-models of prop-wash and scour as a diagnostic evaluation. Although the prop-wash model has been calibrated (in a probabilistic manner) against measurements of near-bottom velocity during ship passage, the

resulting impact on sediment transport has not been calibrated or validated. Impacts include scour and resuspension, and although turbidity data exists that show resuspension events due to propeller-induced scour, these data have not been used to calibrate or validate the model. In addition, the application of the sub-model for propeller-induced scour introduces two new calibration parameters, representing controls on both the erosion as well as deposition of resuspended sediments. This calibration process and calibration parameters represent calibration of both sources and sinks of suspended sediment, potentially resulting in non-unique parameter estimates. It is also not clear what are the reasonable range of values for these new parameters. The future calibration strategy for the propeller-scour sub-model is not clear. The long-term performance of the propeller-scour sub-model is also not demonstrated. The propwash-induced scour can be considered a fully tested and validated sub-model only if shown to suitably reproduce the turbidity (TSS) measurements indicative of the resuspension due to propwash-induced scour and followed by deposition of these sediments. EPA recommends revising the propeller-scour sub-model to (1) avoid additional calibration parameters (this may potentially be achieved by using the measured Sedflume erosion properties and settling velocity established as part of the model calibration), (2) validate the scour and resuspension processes using the turbidity (TSS) signal measured during scour events, and (3) demonstrate model performance over the long-term (the 1999-2012 period used for model calibration).

<u>Specific Comments</u>

1. Section 2.1.3 Hydrodynamic Model, Page 9 First Paragraph: Include the contribution of the tide and estuarine circulation in addition to freshwater inflows from point source discharges in the study area.

2. Section 2.1.4 Sediment Transport Model, Page 12 First Complete Paragraph: Include the contribution of the solids transported by the tide and estuarine circulation in addition to the sediment loadings from point source discharges in the study area.

3. Figure G2-1, Hydrodynamic Model: The graphic only includes flow inputs from point sources and groundwater. Include the tide and estuarine circulation from the East River for completeness.

4. Figure G2-1, Sediment Transport Model: The model framework does not include waves or a bed consolidation algorithm. Although consolidation effects are implicitly included within the model framework by definition of erosion inputs and the fact that depositing sediments recreate the input bed profile of erosion properties, that is not the same as a traditional consolidation model that includes a time- and depth-dependent algorithm of dry density and erosion properties. Remove waves and consolidation from the graphic.

5. Figure G2-1, Sediment Transport Model: The graphic only includes solids loadings from point sources. Include East River solids loadings for completeness.

6. Figure G2-1, Sediment Transport Model: Include settling in the graphic.

7. Section 4.1 Model refinements Made During Phase 2

a. Page 45, There is mention of a radiation separation approach method without any detail or reference. Based on this single sentence, it is difficult to understand how not using this method and applying a new method in Phase 2 helps to improve the model. Provide more information, as appropriate. Also, see Appendix G General Comment #9 in this regard.

8. Section 4.2 Analysis of Phase 1 and Phase 2 Hydrodynamic Data

   a. Page 45, Section 4.2.1 Water Surface Elevation Data. From the Figure, the report claims that only minor differences exist in in tidal amplitude and phase between the two gage locations. Include both in the same figure, and perform and include in revised text a tidal constituent analysis so that amplitudes and phases of the main constituents can be quantitatively compared.

   b. Page 47, Section 4.2.2. Referring to the 10-minute data set, the report mentions the effect of the subtidal oscillation, and the short-duration ebb and flood pulses with relatively large amplitudes during the point source discharge event on July 18. However, there is no mention of the double peaks in ebb and flood currents observed in the 3-hour Low–pass filter time series, and that are caused by the interaction of the New York Harbor and Long Island Sound tides. Expand the discussion of the various features in the data, and the processes/mechanisms responsible for said features.

   c. Page 47, Section 4.2.2 Current Velocity Data. The discussion in this section is focused on 1-week of data that is presented in Figures G4-6 to G4-8, and one single wet weather event. Current profile time-series data was collected for a total of 22 months (Phases 1 and 2) and therefore to limit the discussion on currents to 1 week of depth-averaged currents does not seem appropriate. It is also mentioned that data show a velocity pulse toward the East River and towards land, but there is no explanation of how the discharge generates this sort of back and forth movement of water. Add text/graphics discussing the estuarine circulation process, especially during large point source discharge events and expand the discussion as appropriate.

   d. Page 49, Section 4.2.3 Temperature and Salinity data. Only 3 months of salinity data are available from the continuous time series. The text indicates that based on the discrete salinity data, the overall salinity range is from 1 to 25 PSU. However, table G4-4 shows that the continuous 3-month data only has a range from 6 to 25 PSU, and the majority of the stations do not show values below 10-12 PSU. Revise the report to present an analysis of whether the range of the continuous 3-month salinity data is enough to characterize the conditions in Newtown Creek. It should be noted that only a handful of wet weather events were observed during this 3-month period; these events were also relatively small in terms of the total point source discharge.

      Furthermore, a general summary is presented at the end of this section, but no analysis or detail is provided to support that notion the sets of data are appropriate with respect to having a synoptic understanding of the system using multiple parameters. The main limitation is the short salinity data set, with just a few small point source discharge events. It is doubtful if the salinity data provide enough information to understand the effect of the point source discharges in Newtown Creek for the full range of expected discharge events. Elaborate on these issues in the text.

9. Section 4.2.1 Water Surface Elevation Data, Page 46, 2nd line. The sentence "Tidal motion …" describes a complex tidal regime. This explains why extrapolating the tide from the Battery to the boundary on the other end of the East River is not a good approximation. Use the Lower Passaic River and Newark Bay Superfund or NYC LTCP regional model (or outputs from one of these same models) to specify the tide at the northern boundary; boundary conditions for temperature and salinity can also be specified using the outputs of the selected regional model.

10. Section 4.3 Specifications of Geometry and Bathymetry

    a. Page 50, Section 4.3 The text indicates: "*As such, its boundaries are located 3 to 4 miles upstream and downstream of the mouth of Newtown Creek. It is common practice to set hydrodynamic model boundaries in tidal systems away from the area of interest, to ensure that the numerical methods used to specify inputs at model boundaries do not influence model predictions within the area of interest. That is, establishing the hydrodynamic boundary conditions at locations far from the mouth of the creek was necessary to provide accurate predictions of WSE and current velocity within the Study Area because the parameters are materially affected by circulation patterns and tidal dynamics in the East River*". This sentence provides an explanation of why it is necessary to have the boundaries far enough to provide the correct circulation patterns and tidal dynamics in the East River. Although the locations of the boundaries might be considered far enough from this perspective, if data are not available to create the boundary conditions at one of the selected boundaries, a different location with sufficient data should have been chosen, to guarantee that the model is forced with the correct information. See Appendix G General Comment 1 for corrective actions.

    b. Page 51, Section 4.3. The 2012 bathymetry was averaged into a single cell representative of the average. It cannot be determined if the model resolution is enough that it can maintain geomorphologically distinct features such as the relatively deep navigation channel and sub-tidal flats along the periphery, without losing this in the averaging process. Present a few cross sections in Newton Creek showing how the raw data is represented in the grid.

    c. Page 51, Section 4.3. The report mentions a data gap in bathymetry. Discuss any implication on model results.

    d. Page 51, Section 4.3. The report mentions that near the model boundaries, a constant depth was used to avoid numerical instabilities. Discuss if this is a limitation of the modeling platform, and if it is related to reflection at the boundaries.

11. Section 4.4 Specification of Model Initial and Boundary Conditions

    a. Section 4.4.1. Initial Conditions, Page 52. The report notes that water temperature and salinity were held constant at the initial condition values for the entire 7-day spin-up period. It is unclear how holding the water temperature and salinity constant at both of the East River boundaries for the entire 7-day spin-up achieve a fully "spun-up" condition. It seems like this would generate an artificial condition in which the normal gradients in salinity in the East River were not represented. Normally a hydrodynamic model that is applied to a partially stratified estuary is spun-up (using time varying salinity boundary conditions) for at least one month. Revise the model accordingly.

b. Page 53, Section 4.4.2.1 Water Surface Elevation. The report indicates that "*NOAA tidal gauge data were not available at the northern boundary*". It is correct that WSE data was not available for the full period simulated, but there is WSE data available at Horns Hook (the location of the northern boundary) from 2002 to 2005. These data were also used by NOAA to develop tidal constituents and therefore provide a means to predict the astronomical tide at this location, information that could have been used to generate tidal conditions at the northern boundary instead of a tidal variation based partly on data measured at the Battery. Revise the text to include mention of the WSE data at Horns Hook and why it was not considered for model development.

c. Page 54, Section 4.4.2.1 Water Surface Elevation. Review of the data at the Battery and Horns Hook shows poor correlation between the subtidal fluctuations at these locations. On the other hand, subtidal fluctuations at Horns Hook show a close correlation with the subtidal fluctuations at Kings Point. This indicates that the subtidal fluctuation calculated at the Battery and used to calculate tide at the northern boundary is not correct. Use of a regional model results for tide at the northern boundary will address this issue. See Appendix G General Comment #1 for corrective actions.

d. Section 4.4.2.1 Water Surface Elevation, Pg 54, 3rd paragraph. The report states that application of the Smagorinsky (1963) approach for calculating temporal and spatial variations in horizontal eddy viscosity and diffusivity made it possible to use the tidal harmonic method (the first option) for specifying WSE at the northern boundary and achieve numerical stability. Explain how the application of the Smagorinsky approach "made it possible to use the tidal harmonic method (the first option) for specifying WSE at the northern boundary and achieve numerical stability".

e. Page 54, Section 4.4.2.1 Water Surface Elevation. The text states: "*The amplitude multiplication factors and phase shifts listed in Table G4-11 were adjusted during calibration of the hydrodynamic model as discussed in Section 4.5.1*". The standard practice for numerical model development and application considers model open boundary conditions to be independent of the model calibration process. Various US EPA (US EPA, 2009; US EPA 2010) and International (STOWA/RIZA 1999) guidance documents identifying the individual steps in the life cycle of model development and application consider the specification of boundary conditions a part of the model setup and input. Model calibration is a subsequent and separate process following definition of boundary conditions. See Appendix G General Comment #1 for corrective actions.

f. Page 54, Section 4.4.2.1 Water Surface Elevation. The description of the different options for defining the boundary conditions is not clear in the report. Present more details for defining the boundary conditions to understand the issues of instability mentioned in the report.

g. Page 54, Section 4.4.2.1 Water Surface Elevation. Text states: "*As discussed in Section 4.5.1 the WSE input at the northern boundary was adjusted during model calibration to improve* *prediction of residual flow in the East River*". See comment 11.e above. It is not standard practice to calibrate boundary conditions.

h. Page 55, Section 4.4.2.2 Temperature and Salinity. From the report: "*This assumption is valid because minimal temperature stratification is observed in the East River*." Describe and present what data were used to support this statement.

i. Page 56, Section 4.4.3. Point Source Discharges. The water temperature specified for both discharges from the point source model and the WWTP effluent overflow is the same as for the East River boundary. The text indicates "*This assumption is appropriate because a diagnostic analysis showed that temperature variations in model boundary conditions had minimal effects on hydrodynamic model predictions (see Section 6)*". However, the sensitivity analysis presented in section 6 uses the same temperature values at all the boundaries. A sensitivity analysis that evaluates the effect of the assuming the same temperature for the point sources, WWTP effluents and the East river is not presented. Support the assumption that the temperature of the effluents should be the same as the East River water.

j. Page 57, Section 4.4.4. See comment11; review and address as appropriate.

12. Section 4.5 Calibration Approach and Results

a. Page 59, Section 4.5.1 Calibration Data an Approach. The text mentions the calibration of the boundary: "*The astronomical tide conversion factors used to transform tidal data at the Battery to the northern boundary were adjusted during the calibration process*". See comment 11 regarding the appropriateness of this approach and Appendix G General Comment #1 for corrective actions.

b. Page 60, Section 4.5.1 Calibration Data and Approach. Describe the metrics that were examined in reaching the conclusion that the model is insensitive to effective bed roughness.

c. Page 60, Section 4.5.1 Calibration Data and Approach. The text mentions that the adjustable parameter (AHD) in the Smagorinsky equation is dependent on the spatial resolution of the numerical grid. However, this value has been defined spatially variable from the entrance to the end of the creek, while grid resolution is similar. Explain this inconsistency.

d. Page 62. Section 4.5.3.1 Calibration Results-Water Surface Elevation. The shape of the tide during ebb and flood is not correctly simulated because of the northern boundary. In addition, the subtidal elevation fluctuation at the northern boundary is not correct and can introduce errors. The evaluation of model performance, using bias and ubRMSD, especially for water levels is very limited by methodology. The model could show a small bias error and ubRMSE when long time series are compared (like averaging the error for all conditions), but have significant errors for the conditions that contribute most to the important fate and transport processes in the system. In tidal systems, the assessment of model performance involves examining how the simulated tidal constituents (amplitude and phase) compare to observed values. In addition, performance during events or conditions that are relevant for the project (point source discharge events, surges, etc.)

should also be evaluated. Include and discuss an assessment of model performance by comparing model and data for the amplitude and phase of the major tidal constituents.

e. Page 62. Section 4.5.3.2 Calibration Results-Residual Flow in the East River. This section indicates that the northern boundary condition was adjusted to simulate the average residual flows in the East River (see comment 11 regarding calibration by adjusting the boundary condition). In addition, the target values for the residual flow have a large range. Therefore, the calibration target selected for the model is unclear. It is also not clear why/how important residual flow in the East River is for the project. The report does not present an evaluation of the importance of reproducing the residual flow versus reproducing the instantaneous ebb and flood velocities in the East river which are more relevant to features such as residence time within the model domain. Include a review of instantaneous currents calculated by the model in the East River over a typical spring-neap cycle relative to NOAA measurements.

13. Section 4.5.3.3. Current Velocity

a. Page 65. Section 4.5.3.3.1 Depth-Averaged Current Velocity. There are approximately 21 months of velocity data. The report discusses the evaluation of how the model reproduces the effect of a precipitation event for one case and the report indicates that the model has a relatively good agreement for that event. Expand this discussion to include other conditions, e.g., dry-weather performance, spring-neap performance, etc.

b. Page 65. Section 4.5.3.3.1 Depth-Averaged Current Velocity. The report presents some global results clustering all the data for all the stations, for example saying that for 10 minute results the ubRMSD is approximately 0.1 ft/s. Tables G-17 to G-27 present the ubRMSD by location and deployment. For example, at NC086CM the ubRMSD is approximately 0.22 ft/s and this value is reduced towards land to values of 0.05 ft/s at EK023CM. At the same time that the ubRMSD reduces toward land, the amplitude of the velocities is reduced too. It is important to understand the relative error with respect to the range of values at each location. An ubRMSD of 0.1 might be small at the Newtown Creek entrance, but large towards land. The model/data comparison for currents should include a description of the error at each station including the relative error. Revise the report accordingly.

c. Page 65. Section 4.5.3.3.1 Depth-Averaged Current Velocity. This section does not mention the difference between the simulated and measured currents using the 3-hour low pass filter. The model cannot reproduce the double peak in ebb and flood, which is a consequence of the way the northern boundary has been defined. Revise the text to include a discussion of the features in the data reproduced/not reproduced by the model.

d. Page 66. Section 4.5.3.3.2 Vertical Profile of Current Velocity. Similar to the depth averaged currents, the double peak in ebb and flood is not reproduced by the model . Revise the text to include a discussion of the features in the data reproduced/not reproduced by the model.

e. Page 66. Section 4.5.3.3.2 Vertical Profile of Current Velocity. This section discusses some of the figures in a very general way and some observations from the figures are not mentioned. For example, in Figure G.59 and 60 at NC315 the model seems to overpredict at the surface towards the East River and at the bottom towards land. On the other hand, the text says: "*results indicate that near-surface velocity is overpredicted and near-bottom velocity is underpredicted*". Conclusions of the validity of the results are not made station by station but by averaging and clustering all the stations together. For example: "*the differences in the predicted and observed vertical profiles of velocity are relatively small; on average near-surface velocities are overpredicted by 0.03 ft/s and near-bottom velocities under-predicted by 0.03 ft/s*". These values might look small when multiple stations are lumped together but the conclusions could be different if the error at each station is evaluated relative to the amplitude at that station. Revise the text to include (1) a discussion of the features in the data reproduced/not reproduced by the model, (2) a discussion of model performance (including quantitative comparisons) during dry-weather and large wet-weather periods, and (3) model/data comparison for currents using a description of the error at each station including the relative error.

f. Page 67. Section 4.5.3.3.2 Vertical Profile of Current Velocity. The model performance has not been evaluated independently for wet weather and dry weather. On the contrary, statistics are only presented for the complete time series and in the text for all the stations together. The performance of the model during the wet weather events is very important, and it is important to evaluate the model performance for those specific periods. Include an assessment (both qualitative and quantitative) of model performance separately during dry-weather and wet-weather conditions.

g. Page 67. Section 4.5.3.3.2 Vertical Profile of Current Velocity. The report mentions that the model correctly simulates the temporal variation of the currents during a neap-spring cycle. However, the preceding text does not present a discussion of this feature. Include a description of model performance over the time-scale of a spring-neap cycle.

h. Page 67. Section 4.5.3.3.2 Vertical Profile of Current Velocity. As previously mentioned the parameters used to quantify the model performance (bias and ubRMSD) are calculated for all the depths, all the stations and all the conditions as a single average value. This is not representative of how the model represents different processes. For example, the model could do a good job under normal tidal conditions that are representative of the majority of the time. However, during short-lasting events (e.g., point source discharges, storm surges, etc.), the model may not perform well. In this case, error statistics might be satisfactory, while the model does a poor job reproducing short-term fate and transport which could be important and/or relevant to sediment and contaminant fate and transport. Conduct the model performance evaluation and error calculation with respect to specific processes and types of environmental conditions.

14. Section 4.5.3.4. Temperature

a. Page 67. Section 4.5.3.4 Temperature. It is difficult to conclude from the figures that the larger diurnal temperature fluctuations in the near-surface layer are captured by the model. Revise as appropriate.

b. Page 68. Section 4.5.3.4 Temperature. This is the first reference in the document to a distinction between dry and wet weather conditions in evaluating model performance.

However, there is no explanation on how these conditions have been developed or what they represent. Based on Tables G4-39 through G4-52, wet periods include periods lasting from a few days to almost a month. However, there is no information in the report regarding how these possible periods were selected. Revise the text to include an explanation on how these conditions have been developed or what they represent.

15. Section 4.5.3.5. Salinity

   a. Page 68. Section 4.5.3.5 Salinity. The report states: "*Salinity data collected at the bulkhead sondes are more useful than the salinity downcasts for evaluating model performance because bulkhead sonde data are continuous measurements whereas downcast data are instantaneous measurements*". However, there are only 3 months of data available from the bulkhead sondes and it does not coincide with the period when other variables (water levels, currents, turbidity) were collected. Therefore, most of the model performance analysis and system understanding regarding salinity will have to be performed based on downcast data. Explain if/how the limited amount of salinity data are sufficient for the project.

      i. Page 69. Section 4.5.3.5 Salinity. The available bulkhead data period (July 9 to October 9, 2015) includes two of the driest months of the year, August and September. The model performance during point source discharge events was evaluated for the limited number of events that occurred during this period. The largest events during this period (obtained from figures G-D 146 to G-D-181, note that values in figures G4-81 to G4-86 are in MG/hr) were approximately 70 MG/event, while the annual maximum point source event is in the order of 400 MG/event (from Figures G-D 85 to G-D 134 annual maximum point source event is ~400 MG but it varies from 200 MG for 2012 to 700 MG in 2011). This implies that the 3-month period of available continuous salinity data do not seem appropriate to evaluate the model performance during large wet-weather events. Include a discussion of the available salinity time-series data relative to the environmental conditions in the Creek and whether the salinity time-series data can be considered appropriate for an evaluation of model performance during large wet-weather events.
      ii. Page 69. Section 4.5.3.5 Salinity. Elaborate on why stratification factor was not used for the bulkhead data time series.
      iii. Page 70. Section 4.5.3.5 Salinity. The report states: "*The results discussed above show that the hydrodynamic model simulates salinity with sufficient accuracy to meet the objectives of this study because predicted salinity has minimal bias (typically less than 1 psu) and low ubRMSD (typically less than 1 psu).*" These error statistics, as for other variables, are presented as a global average without separating the performance for different types of conditions that might be more relevant for the project. In this regard the report says: "*model tends to underpredict salinity stratification during wet weather events*". Considering these events are important from the perspective of fate and transport of point source sediment loadings, clarify the importance of the fact that the model does not perform well during these events to meet the objectives of the study. Figures G-D 146 to G-D-181 present the comparison of model predictions to the continuous data from the sondes (July to October 2015). In general, the model consistently underpredicts the variation in surface salinity for most of the point source discharge events, and more clearly for the largest events. These events are important for the transport of point source solids loadings and it is during these

events that the model discrepancies with the observations are the largest. Revise the text to include (1) a discussion of model performance (including quantitative comparisons) during dry-weather and wet-weather periods, and statistical comparison using a description of the error at each station.

iv. Page 71. Section 4.5.3.5 Salinity. The report indicates a number of factors that affect uncertainty in model predictions, but does not present or refer to any analysis that has been done to confirm these as the sources of uncertainty. The horizontal and vertical diffusion were part of the calibration process, probably focused on obtaining the right salinity stratification in Newtown Creek. Elaborate on these factors.

16. Section 4.6 Conclusions

a. Page 71. Section 4.6.1. Overall Hydrodynamic Model Performance. Revise this subsection to clarify the meaning of the following phrase: *"...analysis of predicted WSE versus measured current velocity, salinity and temperature..."*

b. Page 72. Section 4.6.1. Hydrodynamic Model Performance: Water Surface Elevation. As explained in previous comments the predicted WSE are not correct because of the previously described issues with the norther boundary condition for tide. In addition, the text says: "*minimal errors in predicted tidal amplitudes and phase*". Compare tidal constituent amplitudes and phases from the model and the data to evaluate this claim. The text does not present any information besides a few time series plots of water levels that can be used to confirm the claim. See Section 4.5 Comments for corrective actions.

c. Page 72. Section 4.6.2. Hydrodynamic Model Performance: Current Velocity. There is no qualitative or quantitative analysis that demonstrates that the spring-neap variation in currents is well simulated beyond some time series plots, nor does the text include a discussion about it. The estimated errors in current velocities are presented as one single number for all the vertical layers, all the stations and all the periods (dry or wet). The model performance needs to be evaluated by station and with errors relative to the amplitude of the variable at each station and during different periods. A global ubRMSD of 0.15 ft/s seems high – at the mouth of the creek, current amplitudes are in the order of 0.5 ft/s, indicating a relative error of 30%. However, upstream, the amplitudes are much smaller making this ubRMSD value much more concerning. See Section 4.5 Comments for corrective actions.

d. Page 72. Section 4.6.3. Hydrodynamic Model Performance: Temperature. As with other variables only one single value averaged over the whole domain and simulated period is presented to discuss the model performance. The report does not present any analysis of uncertainty during discharge events nor any evaluation to assess if the assumption of using the temperature of the discharge the same as the water temperature at the East River is a valid assumption. The model performance during the specific environmental conditions (e.g., point source discharges, dry-weather conditions, storm surges, etc.) should be assessed in detail. See Section 4.5 Comments for corrective actions.

e. Page 73. Section 4.6.4. Hydrodynamic Model Performance: Salinity. The model underpredicts salinity stratification during wet weather events. These are important periods for point source sediment loadings; however, this is also when the model performance is relatively worse. The lack of continuous data is also a problem for the salinity calibration because the model data comparison is limited to just a few events during the driest months of the year. See Section 4.5 Comments for corrective actions.

17. Section 5.2.1 Multiple Lines-of-Evidence Approach for Evaluating Net Sedimentation Rates

    a. Page 76, Third Bullet in First Paragraph: Sediment traps give information on the gross sedimentation mass flux (in units of mass/area/time), whereas NSRs represent the net sedimentation rate (in unit of length/time). Furthermore, the latter include the effect of spatial variations in dry density in the bed whereas the former do not. Sediment traps are also designed to "trap" suspended sediment that may not be deposited onto the sediment bed. Therefore, due to these reasons sediment trap data cannot be used to develop NSRs as mentioned in the first sentence of this paragraph. They can, however, be used as indicative and qualitative evidence on the sedimentation process, as is described in the fourth paragraph on page 77. Clarify/qualify the use of sediment trap data in the context of the discussion in this section.

    b. Page 76, Fourth Bullet in First Paragraph: Vertical profiles of contaminant concentrations in the sediment bed are mentioned as an approach to develop NSRs. However, subsequent text in Appendix G does not include any mention of this approach. Delete this bullet or add text describing this approach and the results of such analyses.

    c. Page 76, Third Paragraph: As mentioned in the text, although uncertainty in data-based NSRs has been included in the analyses, the potential for bias in any of the individual datasets has not been explored. For example, USACE performance metrics for hydrographic surveys (USACE, 2013) allow for 0.3 ft bias in bathymetric survey data. The resulting error introduced (0.7 cm/yr over 1999-2012, assuming a bias of 0.3 ft in the 1999 data and no bias in 2012) is within the range of sedimentation rates noted in some of the tributaries over this time period (for instance, see the area-Average NSRs in Table G-H-1). One approach to evaluate bias is to compare multiple lines of evidence and check for consistency between the various datasets. In this case, NSRs have been calculated based on geochronology cores, bathymetric differencing over various periods, and historical dredging records. Perform a comparative analysis of NSRs from various approaches and an assessment of the potential for bias in any of the individual NSR approaches. See comments to Attachment G-H for an example of such an analysis for English Kills which indicates a potential bias in the 1999 bathymetry dataset.

    d. Page 76, First and Second Bullets in Fourth Paragraph and associated Figures G5-5 and G5-6: Despite the availability of 1999 bathymetry in Dutch Kills, it has not been referenced in the text or used in the 1999-2012 or 1999-2011 bathymetric differencing. Either (1) include Dutch Kills in these analyses, or (2) provide justification as to why Dutch Kills is being excluded.

    e. Page 76, Third Bullet in Fourth Paragraph: Modify this statement to mention that NSRs were not calculated over the entire area of the East Branch due to partial coverage in 1991.

f. Page 77, Second Paragraph, Bulleted List: Add text discussing the insights regarding historical changes in NSRs and point source sediment loadings resulting from the analysis of NSRs from geochronology cores.

g. Figures G5-5 and G5-6: Either (1) include NSR from historical dredging in English Kills, or (2) provide justification for excluding these NSR estimates.

h. Page 78, Third Bullet in List Continuing from Page 77: Clarify if the temporal variability noted in the gross deposition rates from sediment traps correlate with potential factors such as seasonality in East River TSS concentrations, point source discharge events, storm surges, etc.

i. Page 78, First Complete Paragraph: Some of the insights regarding sediment transport processes have been introduced without presentation of adequate analysis and discussion up to this point in the text. Specific instances are listed below:

   v. The relative distribution of East River and point source loadings
   vi. Impact of propwash resuspension
   vii. Temporal changes in sediment loadings from CSOs

   These are potentially important physical processes at the Site. Provide analyses and discussion to support each of these insights in the various portions of the study area.

18. Section 5.2.2 Data-Based Mass Balance Analysis

a. Page 78, First Paragraph and Equation G-8: There is an *a priori* assumption that no sediment originating from the East River and the Main Stem deposits in the tributaries. This assumption is not discussed in the text. As such, Equation G-8 is missing a term on the right-hand side of the equation representing the net deposition in the tributary of solids originating from the East River and the Main Stem. Either list this assumption and suitable justification, or include the potential for deposition of solids originating from the East River and the Main Stem. The latter alternative can be implemented by replacing term L_{PS} in Equation G-8 with L_{ER+PS}, where L_{ER+PS} represents some unknown combination of solids originating from East River (including Main Stem) and point source loadings. Revise the text accordingly.

b. Page 79, Second Bullet in First Paragraph: Insight regarding the magnitude and composition of point source sediment loadings can be achieved only if assuming no deposition of solids originating from the East River and the Main Stem. If the deposition of solids originating from the East River and the Main Stem is also considered, then no definitive statements can be made on the magnitude and composition of point source loadings. Revise the text by either (1) mentioning that insights about the magnitude and composition of point source sediment loadings can be achieved only under the limiting assumption that no solids from the East River and Main Stem are deposited in the tributaries, or (2) delete this bullet.

c. Page 79, Bulletized list in First Paragraph: While the first bullet is addressed in the results of the analysis (subject to its current assumptions), the goals described in the second and third bullets are not addressed subsequently. Review and revise accordingly.

d. Page 79, Third Paragraph: The last sentence in this paragraph says "*Dutch Kills was not included in this analysis because sufficient bathymetry data were not available*". However, this is contrary to what is described in the following paragraph, that the inputs to this analysis are the USEPA calibration target NSRs. These NSRs are defined in Table G5-8, and include values for Dutch Kills as well. Revise the analysis and text to include Dutch Kills.

e. Table G5-8. Revise the title of third column to "Upper-Bound ...".

f. Page 79, Fourth Paragraph and Figure G5-9: The text and the figure include the statement "*More than 90% of tributary deposition is due to point source sediment loads*". However, there is no text or arguments provided to support and justify this statement. Revise the text and include supporting evidence.

g. Page 80, Equation G-10: The equation for trapping efficiency neglects net deposition in the tributary of solids originating from the East River and the Main Stem. Either list this assumption and suitable justification, or include deposition of solids originating from the East River and the Main Stem. The latter alternative can be implemented by replacing term L_{PS} in Equation G-10 with L_{ER+PS}, where L_{ER+PS} represents some unknown combination of solids originating from East River (including Main Stem) and point source loadings. Revise the text accordingly.

h. Page 80-81, Last Paragraph Starting on Page 80 and Figure G5-13: The calculations presented in this section assume that point sources are the sole source of depositing sediments to the tributaries. This is an unsupported assumption. For the example of English Kills presented in Figure G5-13, using average tidal range of 1.5 m, area of 94,500 m^2, and a nominal 10 mg/L of TSS gives gross annual solids load of ~1000 MT/yr imported from the Main Stem during the flood phase of the tide (with unknown export during the ebb phase of the tide), a number in excess of even the upper uncertainty bound (910 MT/yr) in Figure G5-13. The potential for deposition of this load from downstream (from the main stem) is not considered in the mass balance calculations. Rather, the statement "*Valid Assumption: Sediment loads from downstream sources have relatively minor effect*" is made in Figure G5-13 without any supporting evidence. Either list this assumption and suitable justification, or consider the potential for deposition of solids originating from the East River and the Main Stem. Revise the text accordingly.

i. Page 80-81, Last Paragraph Starting on Page 80, continuing to Page 81, and Bullet List on Page 81: The results of the sediment mass balance analysis do not seem to be referenced anywhere else in the text. How have the results of this analysis been used subsequently? Either (1) refer to this analysis in a following section, or (2) delete this section.

19. Section 5.2.3 Bed Property Data

a. Pg 81. Revise the report to describe what causes the bed composition to become coarser moving upstream from the East River (it is also generally coarser in the tributaries)

b. Page 81, First Complete Paragraph: The reference to fluid mud is made rather abruptly at the end of the paragraph and without any context. Is the assertion that fluid mud is present in areas upstream of CM 1 where on average, dry density is less than 0.4 gm/cm^3? Clarify the text.

c. Page 81, Second Paragraph: There is large variability in the fines content within the main stem and the tributaries. For instance, as seen in Figure G5-22, fines content ranges from ~15-100% between CM 0-1. Is this spatial heterogeneity related to features such as point source release location or other factors such as the flow characteristics of the water body? In relative terms, point source loadings are comprised of more sands (~40-50% as per Table G5-6) than East River loadings (2% as per Section 5.4.1). Assuming that sands are deposited in the proximity of the outfalls, this could potentially explain the spatial variability in fines content. The spatial heterogeneity of fines content could be relevant to the contaminant fate and transport modeling efforts since contaminants typically partition to organic carbon-rich fine sediments more than sands. If so, it may be of use in refining the model initial conditions. Review the data and clarify if the heterogeneity can be explained by afore-mentioned factors, and incorporate into the model as appropriate.

20. Section 5.2.4 TSS Concentration and Turbidity Data

a. Page 82, First Paragraph: Clarify the conclusion from Figure G5-28 – is there or is there no temporal trend in TSS at the mouth of the Creek?

b. Page 82, Second Paragraph: See comments on Attachment G-F. There is a correlation between turbidity and TSS, primarily dependent on environmental conditions (dry-weather versus large wet-weather events). The resulting turbidity-TSS relationships can be used to develop TSS time-series. The TSS time-series can be used to calibrate the sediment transport model during dry-weather and large wet-weather event conditions which represents two bounding conditions for sediment transport. Revise the text and figures to (1) include a discussion of the turbidity-TSS correlations, (2) develop estimates of TSS time-series from the measured turbidity, (3) use the resulting TSS time-series in developing an understanding of sediment transport within Newtown Creek (for instance, dry-weather versus wet-weather conditions), and (4) use the TSS time-series estimates as a model calibration metric.

21. Section 5.3.1 Sediment Size Class Characteristics

a. Page 84, First Paragraph: The choice of the number and type of sediment size classes (how many cohesive classes, and how many non-cohesive classes), is typically made based on site-specific factors such as the sediment substrate, analysis of TSS time-series data, etc. However, the text does not currently provide such explanation. Explain the rationale and provide evidence supporting the choice of sediment classes included in the model.

b. Page 84, First Paragraph: A row of cells across the mouth of Newtown Creek (I=12) appears to have been defined as hard-bottom even though these seem to be partly within the boundary of the Study Area. Review and revise as appropriate.

c. Page 84, Third Paragraph: The selection of particle diameters for class 2 and class 3 (fine sand and medium-coarse sand, respectively) seems to have followed different procedures. Class 2 particle diameter was determined based on an assumed settling velocity. In contrast, Class 3 particle diameter was first assumed and a corresponding settling velocity calculated. However, neither of these particle diameters are data-based, i.e., based on an analysis of the grain size distribution within Newtown Creek. Given the relevance of particle diameters on the erosion, armoring, and hiding/exposure functions inherent in the active-layer formulations of SEDZLJ as well as settling velocity, the particle diameter inputs should be based on an analysis of bed grain size distribution measured within Newtown Creek. Particle diameter is the fundamental sediment property from which other characteristics such as settling velocity and erosion-behavior (via the critical shear stress for erosion, armoring, and hiding/exposure, etc.) are derived. There are many methods for calculating a representative particle diameter for given size class. One approach would be to calculate the median diameter within a size class (e.g. between 63 -250 um) for a given core, and then calculate an average diameter for all the cores within Newtown Creek and use for model input. Finally, settling velocity should be calculated based on particle diameter, not the other way around. Revise the model inputs and text accordingly.

22. Section 5.3.3.1 East River

    a. Page 88, Last Paragraph: Clarify what the conclusion is from the temporal trends in TSS shown in Figure G5-36. The data seem to indicate a seasonal trend, declining through summer and fall before increasing in the winter and spring, a seasonality similar to freshwater flow in the Hudson River.

    b. Page 89, First Sentence: A vertically constant profile of TSS was applied at the East River boundaries. Do the data used to develop the boundary conditions (data near the mouth of Newtown Creek) show any vertical gradients in TSS? Such gradients are typical of fine sediments, and in combination with estuarine circulation can result in net upstream flux of fine sediments. If the data show vertical gradients, then apply such a gradient the East River boundaries.

23. Section 5.3.3.2 Point Source Discharges

    a. Page 90, Last Paragraph: Relate the analysis described in this paragraph to the remainder of the text in this section.

    b. Page 91, Third Paragraph: Explain why no wash-load fractions are assumed to be associated with point source discharges.

24. Section 5.4.1 Calibration and Validation Approach

    a. Page 94, First Complete Paragraph and Figure G5-45: Reconcile the text in this paragraph with Figure G5-45. The text indicates that NSRs were the only calibration target, with the other metrics listed in Table G5-7 used for model validation. Figure G5-45 makes no such distinction; instead it indicates that all the metrics listed in Table G5-7 were used for model calibration.

    b. Table G5-8: The third column is mislabeled as "*<u>Lower</u>-Bound…*"; it should be "*<u>Upper</u>-Bound…*". Revise accordingly.

c. Page 94, Second Paragraph, Figure G5-46, and Table G5-9: Explain the rationale behind the choice of NSR calibration targets using different approaches in various portions of the study area. The NSR calibration targets appear to have been defined using a number of somewhat inconsistent approaches. For instance, NSR calibration targets in the main stem were defined using 1991-2012 bathymetric differencing, East Branch and Maspeth Creek using 1999-2012 bathymetric differencing, and English Kills using the lower bound of USEPA-proposed NSR ranges despite the availability of 1999-2012 bathymetric differencing. In addition, explain why a NSR calibration target is not defined for Dutch Kills despite the availability of 1999 bathymetry data as well as USEPA-proposed NSR ranges.

d. Page 94, Second Paragraph, Figure G5-46, and Table G5-9: As mentioned in the comments to Attachment G-H, the 1999 bathymetry data may likely be biased in English Kills and the East Branch and therefore unsuitable to establish NSR calibration metrics. In contrast, the 1991-2012 NSR is consistent with the other lines of evidence for NSRs in these tributaries. Within Maspeth Creek, barring an area of high sedimentation near the mouth (Area 2 in Table G-H-3), the 1991-2012 and 1999-2012 bathymetric differencing produce relatively similar NSRs as other lines of evidence. Therefore, it is appropriate to use the 1991-2012 bathymetric differencing to define calibration targets in Maspeth Creek (primarily for portions away from the mouth; NSRs shown in Table G-H-3), resulting in calibration targets within the USEPA NSR ranges for this tributary. Finally, in Dutch Kills, which was not covered in the 1991 survey, the 1999-2012 bathymetric differencing indicates NSRs relatively similar to other lines of evidence (adjusted for uncertainty in Pb-210 NSRs). However, the NSR calibration target calculated using the 1999-2012 bathymetric differencing is higher than the upper bound of the USEPA NSR ranges for this tributary. Figure 1 shows a graphical comparison of the USEPA NSR ranges, the FMRM NSR calibration targets, and proposed EPA NSR calibration targets. The proposed revisions to the NSR calibration targets also represents a more consistent use of datasets than the approach in the FMRM which uses 1999-2012 bathymetric differencing in two tributaries, the lower bound from USEPA's NSR ranges in another tributary, and no calibration target in the fourth tributary. The proposed approach relies on the 1991-2012 bathymetric differencing in three tributaries, using the 1999-2012 bathymetric differencing in the fourth tributary solely due to a lack of bathymetric coverage in 1991. Furthermore, the proposed NSR calibration targets are also consistent with the NSRs from other lines of evidence. Perform a comparative analysis of NSRs from multiple lines of evidence as a data quality check on the NSRs from individual approaches, and develop a consistent approach for defining NSR calibration targets in the various tributaries.

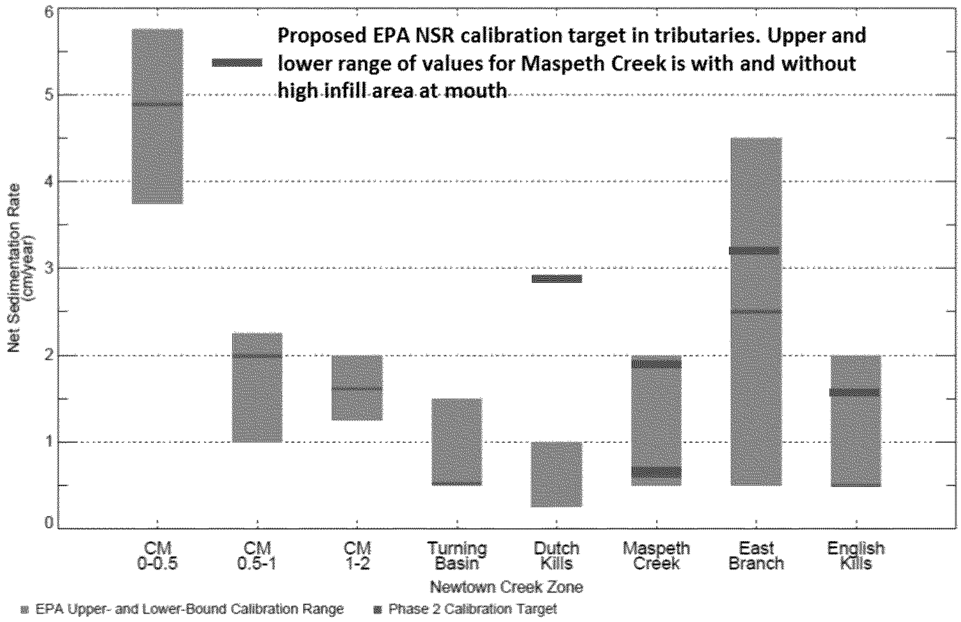

Figure 1 USEPA NSR ranges, FMRM NSR calibration targets, and proposed NSR calibration targets. Note, proposed EPA NSR calibration targets may differ from NCG analysis presented in Attachment G-H due to differences in analytical methodology.

This analysis for reconciling NSRs from various lines of evidence suggests that the 1991-2012 bathymetric differencing is appropriate to define NSR calibration targets in English Kills, East Branch, and Maspeth Creek. The resulting NSR calibration targets are also within the USEPA NSR ranges for these tributaries. Within Dutch Kills, due to the lack of bathymetry data from 1991 and because the 1999-2012 NSRs are consistent with other lines of evidence, the 1999-2012 bathymetric differencing is appropriate to define NSR calibration targets. However, the resulting NSR calibration target is higher than the upper bound in USEPA's NSR range for this tributary. The proposed revisions to the NSR calibration targets for these tributaries also represents a more consistent use of datasets than the approach in the FMRM which uses 1999-2012 bathymetric differencing in two tributaries, the lower bound from USEPA's NSR ranges in another tributary, and no calibration target in the fourth tributary. The proposed approach relies on the 1991-2012 bathymetric differencing in three tributaries, using the 1999-2012 bathymetric differencing in the fourth tributary solely due to a lack of bathymetric coverage in 1991.

e. Page 94, Third and Fourth Paragraphs: A number of inputs and parameters were adjusted as part of model calibration:

   i. Model inputs
      1) East River wash load content
      2) East River flocculated clays/silt content
      3) East River fine sand content
   ii. Model parameters
   iii. East River flocculated clays/silt settling velocity

iv. Point source flocculated clays/silt settling velocity

In addition, the settling velocity of wash load is an assumed value. Because it is not based on site-specific data, it is not a truly independent parameter. In other words, a different settling velocity assumption could require a different East River wash load content to reproduce the performance obtained with the FMRM parameterization. A similar argument exists with respect to the settling velocity and East River content for the flocculated clays/silts as well, where an increase in settling velocity could potentially be compensated by a decrease in East River content for this size class. In other words, the large number of inter-dependent model assumptions, and model inputs/parameters subject to calibration indicates the potential for non-unique input and parameter combinations which in turn reduces confidence in model predictability and performance. Develop an approach that can help reduce the number of model inputs and parameters subject to assumption and/or calibration.

f. Page 94, Third and Fourth Paragraphs: Inputs such as the mass fractions of the three size classes in East River suspended sediments, essentially the boundary conditions, should not be subject to calibration. According to guidance from USEPA (2009, 2010) and others (STOWA/RIZA, 1999), in the process cycle of model application for a given site, model inputs such as boundary conditions should be defined separately from and prior to the process of model calibration. The process of model calibration should focus on model parameters such as settling velocity rather than boundary conditions. Therefore, model inputs such as the mass fractions in East River loadings should be determined either on the basis of measurements, or on the basis of suitable data analysis. Develop data-based and/or empirical approaches to constrain the sediment mass fractions in the East River loadings. See the following comment for additional suggestions in this regard.

g. Page 94, Third and Fourth Paragraphs: EPA recommends measurements of grain size distribution in the East River loadings (perhaps as part of future sampling) which will help constrain these model inputs. In the interim, given the lack of such data, there are potential analytical approaches that may help estimate the composition of East River loadings. The sand content of East River loadings may potentially be calculated based on a sediment mass balance. Assuming sands transported from the East River are deposited within CM 0-2 (this length corresponds to the tidal excursion length for a particle located at the mouth of Newtown Creek at the beginning of the spring flood tide), the likely sand loading from East River can be calculated as follows:

East River sand content (mass/volume) = [(measured sand content in the sediment bed between CM 0-2 * dry density * NSR * bed area) – (estimated annual point source sand loadings between CM 0-2)] / [Tidal Prism of entire Newtown Creek at average tidal range * Number of tides/yr]

The relative distribution of the flocculated clays/silts and wash load fractions can potentially be determined by reviewing available data – for instance, as shown in Figure 2, median depth-average TSS concentrations during the flood and ebb phases of the tide correspond to approximately 30 mg/L and 20 mg/L, respectively, in the vicinity of CM 0.25.

This relative difference between flood and ebb provides an approximate idea of the relative proportions of the fast-settling and slow-settling classes.

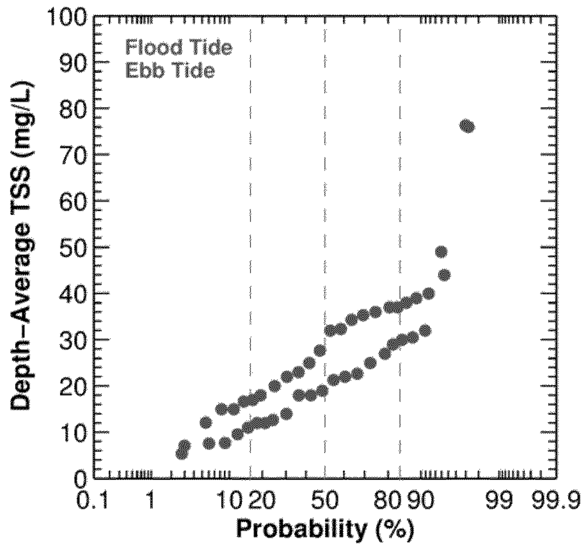

Figure 2. Probabilistic comparison of depth-average TSS measurements during flood and during ebb using measurements during Phase 1 and Phase 2. Data from the vicinity of CM ~0.25

h. Page 94, Fourth Paragraph: Elaborate on the rationale and basis for the difference in settling velocity of flocculated clays/silts from East River and from point sources, at 3 and 1 m/d, respectively.

i. Page 94, Fourth Paragraph: The calibrated settling velocity for the flocculated clays/silts from East River and from point sources at 3 and 1 m/d are about 10X too low compared to estimates from within Newtown Creek and compared to other studies of NY harbor (LBG et al., 2014; HydroQual, 2007; Ralston et al., 2013; Fugate and Chant, 2006). In particular, Fugate and Chant (2006), sampling the CSO plume from an outfall in Flushing Bay, NY, estimated settling velocity for CSO solids ranging from 43 m/d to 800 m/d, with a median value of 250 m/d. The sediment transport model developed by Moffatt & Nichol and Deltares for the Lower Passaic River and Newark Bay Superfund sites (currently under final review by US EPA Region 2) also uses fine sediment settling velocities up to an order of magnitude higher than used in the FMRM. In addition, site-specific estimates within Newtown Creek have been derived using the gross sedimentation rates measured in the sediment traps (data shown in Figure G5-7). These data were paired with the fines content measured for the sediment accumulated in these traps, median near-bottom TSS from Phase 1 and Phase 2 measurements in the vicinity of the traps, and an average spring-neap probability of deposition at the trap locations (assuming critical shear stress for deposition of 1 dyne/cm^2 and the Krone formulation for probability of deposition). The calculations were performed for various locations along the main stem and a location within lower English Kills. Since the sediment traps measure gross sedimentation rates, the settling velocity estimated using this approach is an estimate of the gross settling velocity, a number directly comparable to model inputs for this parameter. Figure 3 shows the results of this analysis in comparison to model inputs (horizontal dashed lines). The comparison shows that barring a few instances in the vicinity of the Turning Basin, the majority of the estimated settling velocity values are higher (up to 10X) than model inputs for flocculated fines from East River as well as point sources.

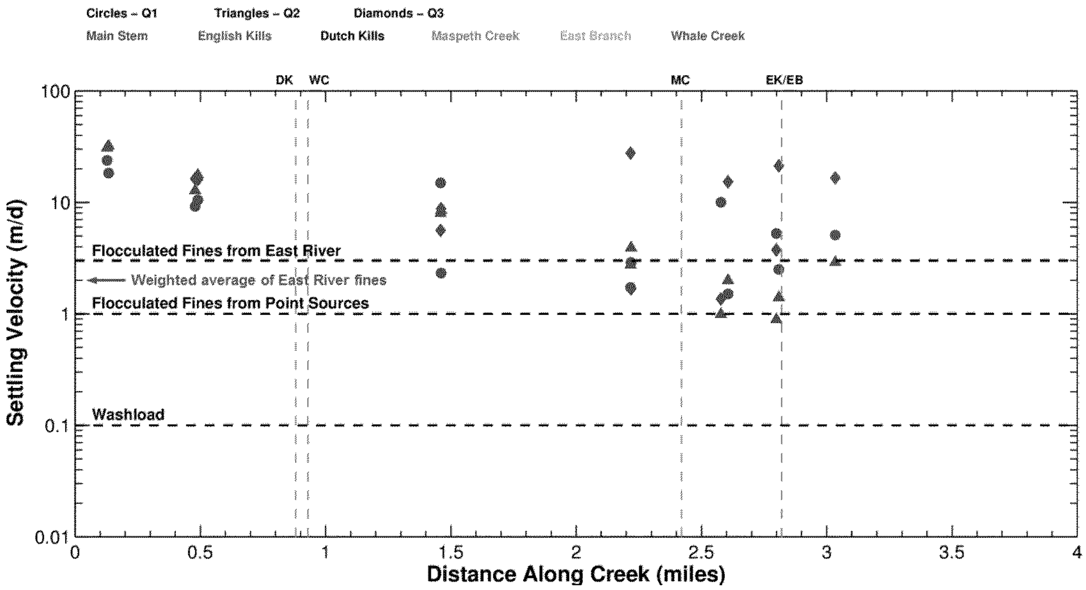

Figure 3. Spatial profile of settling velocity estimated from sediment trap data at selected locations relative to model inputs for settling velocity.

Reconcile model inputs for settling velocities with these site-specific estimates.

25. Section 5.4.2.1 Model Calibration: NSRs for 1999 to 2012

   a. Page 95, Third Paragraph: In addition to the area-average NSR, include a probabilistic comparison of NSRs from bathymetric differencing averaged over the scale of model grid cells, and model-calculated NSRs. The comparisons can be over the same reaches as used in Figures G5-47 through G5-49. This will allow for an assessment of how well the model captures the spatial variations in NSRs within individual reaches.

   b. Pg 95, Last paragraph. Deviations between predicted and data-based NSRs are most likely due to more than just the "uncertainty in the magnitude and composition of point source sediment loads for these two tributaries". Revise the report to include additional evaluations of the source of the observed deviations between predicted and data based NSRs

   c. Page 96, First Paragraph: Include a spatial (map) comparison of NSRs over the scale of individual model cells (using 1991-2012, and 1999-2012 bathymetric differencing, as appropriate) and model results. This will allow for an assessment of the spatial pattern of NSRs and how well the model performs relative to data.

   d. Page 96, Second Paragraph: What is purpose behind comparing parallel and continuous simulations? Will parallel simulations be performed in the future or is the continuous simulation approach the preferred approach for the FS simulations? Provide clarifying text to address the questions and, if the latter, delete this paragraph and associated figure.

26. Section 5.4.2.2 Model Calibration: NSRs for 1999 to 2012

    a. Page 96, Fourth Paragraph: In addition to the area-average fines content shown in Figures G5-52 to G5-54, include a probabilistic comparison of the fines content in individual surficial cores, and model-calculated fines content in individual grid cells. The comparisons can be over the same reaches as used in Figures G5-52 to G5-54. This comparison allows for an assessment of how well the model captures the spatial variations in fines content within individual reaches. Review of the FMRM results in this fashion shows spatial patterns in the model-calculated fines content and deviations from data distributions which may be indicative of sand loadings from point source discharges and fine sediment transport from the main stem into some of the tributaries. Review and elaborate as appropriate.

    b. Figure G5-54: There appears to be a minor bug in the model outputs, affecting the model results for CM 0-0.5 shown in Figure G5-54. Reviewing the 15 cm composition output by the model in file Graphics_bin.out, for a row of cells across the mouth of Newtown Creek (I=12), the fines content at the end of the simulation is reported as 0 even though this row of cells doesn't see any erosion or deposition and so shouldn't deviate from the initial condition of ~90% fines. The area-average fines content for CM 0-0.5 using the 15 cm composition output results in ~50% fines, as shown in Figure G5-54. However, using the fines content in the top 15 cm of the bed calculated from the model restart file at the end of 2012 results in ~70% fines content for this area, a number more similar to the data. Review and address as appropriate.

27. Section 5.4.2.3 Model Validation: TSS Concentration for 2012 to 2015

    a. Page 97, First Paragraph: The TSS model-data comparisons have been presented in terms of individual spatial profiles. However, this prevents an objective assessment of model performance across the entire dataset. This can be achieved using cross-plots of model-calculated TSS versus measured TSS, and probability plots of model-calculated TSS and measured TSS. Review of the FMRM model performance in this fashion shows a bias towards under-prediction with distance upstream in Newtown Creek. This implies that the model likely does not capture the gross tidal transports into and out of the tributaries and in the main stem upstream of CM ~2. In addition, review of the Phase 1 and Phase 2 TSS data show specific trends such as higher concentrations during the flood phase than during ebb phase of the tide, and higher concentration during spring tides than during neap tides. Both trends are physically reasonable and explainable, and are true for most locations within the study area. However, these trends are not reproduced by the model. Model performance for TSS should be considered as a calibration metric rather than as part of validation. This is also relevant for the contaminant fate and transport and food chain models, since TSS concentrations could control contaminant particulate-phase concentrations, and resulting food chain exposure concentrations. Include model-data comparisons for TSS (from water samples as well as estimates from turbidity measurements) as a calibration metric.

b. Page 97, Last Paragraph, First Bullet: Examine the East River TSS data for seasonal trends and incorporate in the model as appropriate. As mentioned in the comments to Section 5.3.3.1, there may be a seasonal trend apparent in the data. In addition, review the East River TSS data for vertical gradients and incorporate in the model as appropriate. As mentioned in the comments to Section 5.3.3.1, the vertical gradient in TSS in combination with estuarine circulation is a process that can potentially result in net upstream transport of fine sediments.

c. Page 98, Paragraph Continued from Page 97, Last Bullet: The impact of neglecting primary production of solids can be assessed by reviewing the model-data comparisons on a seasonal basis, separately for May-September and October-April. The former corresponds to the period expected to be affected by primary production and vice versa for the latter. Perform model-data comparisons on a seasonal basis to evaluate the potential for primary production to bias the model-data comparisons. If the wintertime model-data comparisons are similar to summertime model-data comparisons, then there is no likelihood of primary production introducing a bias in model-data comparisons. If this is true, delete this bullet.

28. Section 5.5.1.1 Diagnostic Analysis: Continuous versus Superposed Simulations

    a. Page 98, Third Paragraph: As mentioned in the comments to Section 5.4.2.1, the purpose of this analysis is not apparent. Will parallel simulations be performed in the future or is the continuous simulation approach the preferred approach for the FS simulations? Revise the text to address the purpose of this analysis or, if the latter, delete this paragraph and associated figure.

29. Section 5.5.1.2 Diagnostic Analysis: Relative Effects of East River and Point Source Sediment Loads

    a. Page 99, Paragraph Continued from Page 98 and Figure G5-60: Figure G5-60 is a nice presentation of model performance. Add a similar figure in Section 5.4.2.1 along with another line to indicate the measured laterally averaged NSR. This will allow another type of assessment of model performance relative to data.

    b. Page 99, First Complete Paragraph: Add text with rationale for why 2009 was selected for this diagnostic simulation and if 2009 is a typical year with respect to point source loadings.

30. Section 5.5.1.3 Diagnostic Analysis: Sediment Mass Balances, Page 101, 3rd bullet. Revise the report to describe the impact of the assumed constant SSC boundary conditions at the East River boundaries on the 4,900 MT/year net incoming sediment load.

31. Section 5.5.3 Diagnostic Analysis of Direct Geomorphic Feedback, Pg 104. Revise the report to explain how "direct feedback between the hydrodynamic and sediment transport models" was accomplished.

32. Section 5.5.4 Diagnostic Analysis: Sediment Mass Balances

    a. Page 100, Equation G-18: This equation neglects any import from point sources elsewhere in the domain and from the East River loadings. In other words, trapping efficiency for a given tributary is calculated relative only to the point sources loading in that tributary. It ignores other sources of sediment loading to the tributary such as sediment discharged from point sources elsewhere in the domain and sediment from the East River loadings. Either (1) list this assumption, or (2) revise Equation G-18 appropriately, considering all sources of sediment loadings.

b. Page 101, Bullet List carried over from Page 100 and Figure G5-65: There seems to be an inconsistency between the information in the bullet list, and the model results shown in Figure G5-63. The trapping efficiency for English Kills is listed as 100% in Figure G5-65 – 230 MT/yr of sediment is discharged from the point sources in English Kills, and 230 MT/yr of deposition is shown in English Kills, which implies that all the sediment discharged from the point sources in English Kills is trapped within this tributary. However, Figure G5-63 shows that ~5% of the sediment deposited within English Kills originates from the East River. In other words, ~12 MT/yr of the 230 MT/yr represents East River solids. This is inconsistent with the 0 MT/yr exchange between English Kills and Main Stem indicated in Figure G5-65. In addition, trapping efficiency, as written in Equation G-18, can only be a maximum of ~95%. Furthermore, of the ~218 MT/yr of point source loadings depositing in English Kills, it is not clear if any of this sediment originates from point source releases from elsewhere in the Study Area, e.g. East Branch, Turning Basin, etc. Accounting for all the sediment loadings to a given reach (point sources within reach, and advection from downstream) in Equation G5-18 will address this issue. Review this issue and address in the text and figures as appropriate for the other reaches and tributaries listed in the text in this section and in Figures G5-65 to G5-71.

33. Section 5.5.4 Diagnostic Analysis of Organic Carbon Solids Transport

a. Page 106, Second Paragraph: The following statement is made regarding the type of organic carbon (OC): "*The data-based results discussed above show that TOC content in bed sediment in the tributaries (approximately 10 to 20%) and f_{OC} in CSO and stormwater discharges (average of 16%) are similar. This consistency between f_{OC} in point source discharges and TOC content in bed sediment indicates that OC solids in point source discharges are primarily composed of G3 OC, with relatively minor amounts of G1 and G2 OC. Thus, OC solids in the sediment transport model diagnostic simulation were represented by assuming that 100% of the OC solids were the very slowly decaying (G3) OC fraction.*"

The above rationale for making this conclusion cannot be justified. Just because the f_{OC} values for CSO, stormwater, and point source discharges are similar to f_{OC} of tributary bed sediment does not justify assuming that all OC loadings are G3. Non-point and point source loadings are most likely composed of both labile and refractory OC. The G fractions are typically used to distinguish the benthic sediment OC types, not the OC of the water column and loadings. The usual practice is to assign labile OC deposited to benthic sediment to the G1 class and to split the refractory OC deposited to benthic sediment between the G2 and G3 classes. It is highly doubtful that the OC loadings from non-point and point source loadings are all highly refractory and associated with the G3 class following deposition. CSO and wastewater treatment plant loads are likely to have considerable labile OC which would fit into the G1 class when deposited to benthic sediments. Therefore, it is unreasonable to assume that all point source OC loadings are highly refractory (i.e. G3). Point source OC loadings comprise a mix of labile and refractory forms of OC. Following deposition to the bed, the G1 (labile) and G2 forms of OC will degrade and cannot be assumed to be conservative. In addition, primary production in the water column may also provide an additional source of OC which may subsequently be deposited to the bed. Revise this diagnostic analysis by considering the various forms of OC appropriately – primary production, labile and refractory in the water column, and G1/G2/G3 in the sediment bed.

b. Page 106, Third Paragraph: The model OC is distributed among four size classes that correspond to sediment size classes in terms of settling rates (see Table G5-13). There is no explanation of how the OC is split among these four size classes. Add a discussion and justification to explain how the OC is fractioned.

c. Page 106, Third Paragraph: The length of the diagnostic simulation was one year, which is not long enough to properly evaluate the adequacy of the organic solids transport model. The net sediment rate is on the order of ~1-2 cm/yr, so for a one year simulation, the depositional contribution to the sediment bed is small relative to the mass within the model bed layer. Thus, with such a short simulation period, the model results at the end of one year will be very similar to the initial conditions. It then becomes relatively easy to force model agreement with observations by adjusting the initial conditions. Revise this diagnostic analysis with a much longer simulation period (over the 1999-2012 period used for the sediment transport model) for proper diagnostic evaluation.

c. Page 107, First Complete Paragraph: Results of a one-year diagnostic simulation are compared with observed surface sediment TOC concentrations in Figures G5-90 – G5-93. Recognizing the shortcomings of the short simulation period where one-year results are similar to initial conditions (preceding comment), the model results compare relatively well with observed TOC along Newtown Creek (Figure G5-90). The spatial averages for model results and observed data along Newtown Creek also compare favorably (Figures G5-91 – G5-93). However, there are no spatial comparisons of computed and observed data within East Branch, Maspeth Creek, and Dutch Kills. Present a comparison of model results along each tributary reach with observed data from the tributaries, without spatial averaging, to show how well the model performs and to determine if the model satisfactorily exhibits the rather rapid upstream increase in tributary bed TOC (i.e., 10 to 20 % TOC).

d. There are no comparisons of model results with observed suspended OC data (particulate organic carbon, POC) for the water column. Such comparisons are necessary to obtain a complete picture of model performance. Revise the document to include such comparisons.

34. Section 5.5.5 Diagnostic Analysis of Hard Bottom Assumption in East River

a. Page 108, First paragraph. The report states that the only source of sediment that was transported into and out of the active surface layer was suspended sediment in the East River. Revise the report to identify other sediment sources within the system.

b. Page 108, Second Paragraph: The diagnostic simulation shows higher net solids flux from the East River into Newtown Creek. However, the text does not explain this result and the transport mechanisms responsible for this result. Elaborate upon this result in the text.

c. Page 108, Second Paragraph: Elimination of the hard-bottom assumption in the East River likely leads to erosion and deposition over tidal time-scales within the East River. Since this is a realistic phenomenon for such tidal systems, there is a physical basis and argument for not including a hard-bottom assumption in the East River. Eliminate the hard-bottom assumptions in the East River.

35. Section 5.5.6 Diagnostic Analysis of Propwash Resuspension

    a. Page 109, First Paragraph: Elaborate on why propwash-induced scour is important in Newtown Creek. Besides the existence of localized scour holes as evident in the multi-beam bathymetry data, what other evidence exists that provides an idea of the relative importance of propwash-induced scour relative to normal hydrodynamic forcings (tides, point source discharges, estuarine circulation, etc.)? In other words, how important is propwash-induced scour to the large-scale spatial and temporal patterns of suspended sediment transport? Revise the text accordingly.

    b. Page 109, Last Paragraph, Fourth Bullet: Elaborate on why the direction of transit (inbound or outbound) matters for propwash and scour.

36. Section 5.5.6.1.2 AIS Data Analysis: Historical Data

    a. Page 112, Second Paragraph: Define what is indicated by the term "Ship days" which first appears in Figure G5-103. Also, elaborate on how the information in Figure G5-103 does not represent a complete picture of navigation traffic because of the discrete nature of AIS data.

37. Section 5.5.6.3.2 Development and Calibration of Empirical Propwash Model

    a. Page 118, Second Paragraph: Revise the text to indicate that the AIS data does not provide information on the actual draft which depends on whether the vessels are loaded or not. Rather, AIS data only provides the rated draft which (in combination with the local instantaneous water depth) does not provide a true measure of the distance between the propeller shaft and the sediment bed.

    b. Page 118, Third Paragraph Bullet List: Add uncertainty on the actual vessel draft to this list.

38. Section 5.5.6.4.1 1-Year Diagnostic Simulation: Single Representative Ship

    a. Page 120, Last Paragraph Bullet List: The two potential calibration terms listed here represent a control on the erosion (first bullet, Probability of resuspension), and a control on deposition (second bullet, Effective settling speed of resuspended Class 1 sediment). Calibrating both the erosion and deposition process in this context can lead to non-unique solutions. For instance, a given TSS response in the water column can be achieved as the net of two large parameter values for the two terms, or as the net of one moderately-high term (for example, the erosion-related term) and a deposition term with relatively average value. Develop an approach that minimizes the need for calibration. Also, see the next comment.

    b. Page 120, Last Paragraph First Bullet: It is not clear why the erosion due to propwash needs to be subject to calibration. The Sedflume data summarized in Section 5.3.2 in combination with the site-specific data on particle diameters, and grain size distribution should in principle be adequate to characterize the erosion properties of the bed due to typical hydrodynamic forcings (tides, point source discharges, etc.), and propeller wash. Given the existing erosion parameterization and the context of the preceding comment, do not use the probability of resuspension due to propwash as a calibration parameter. Revise accordingly.

   c. Page 120, Last Paragraph Second Bullet: Provide sufficient justification and evidence why the settling speed of Class 1 sediment resuspended due to propwash scour should be different from the settling velocities used for Class 1a and Class 1b in the base calibration simulations.

39. Section 5.5.6.4.2 1-Year Diagnostic Simulation: Multiple Ships

   a. Page 123, First Paragraph: Why does the ship traffic for 2009 (Figure G5-141) look dramatically different from 2010 (Figure G5-132) upstream of CM 1? For instance, Lower English Kills sees 300-400 Ship-days of vessel traffic in 2010 but only 1-50 Ship-days in 2009? Traffic seems to have increased by a factor of ~10 in a 1-year period. Revise the text accordingly.

   b. Page 124, First Paragraph First Bullet: How were the simulations with propwash scour judged to "yield realistic predictions"? Describe in detail the process whereby model performance with propwash scour is assessed and judged, describe the model calibration process, and present a comparison of model and data.

   c. Page 124, First Paragraph First Bullet: Realistic predictions of propwash scour are mentioned as being generated by "adjusting input parameters within the range of values used in the diagnostic analysis". Describe what a realistic range of parameter values should be for the two calibration terms related to propwash scour predictions. For instance, the 200 m/d settling velocity used in one of the diagnostic simulations corresponds to nearly fine sand, whereas the surficial sediments in the majority of the Study Area are known to be comprised of fine sediments.

   d. Page 124, 2nd bullet. The report states that "Predicted NSRs are … relatively insensitive to variations in effective particle diameter, when the settling speed was at the upper-bound value (200 meters/day)". The significance of this statement is unclear. Revise the report to describe why there are differences in the predicted NSRs with particle diameter if the settling speed was held constant.

40. Section 5.5.6.4.3 Path Forward

   a. Page 124, Second Paragraph: The propwash resuspension model is described as producing "*realistic results that are qualitatively representative*". However, the preceding text does not discuss any of the four diagnostic simulations with varying parameter values for the two calibration parameters in the propwash scour model. It is not clear what the proposed parameter value is for either of these terms. Nor is it clear what rationale was used to judge the model performance as realistic and qualitatively representative. Add sufficient discussion of the model results and arguments leading to these conclusions.

41. Section 5.6 Conclusions

  a. Page 124, First Paragraph: The first sentence and the bullet list is somewhat confusing. Is the intent to list model inputs for which adequate data was available and used to specify inputs, or is it meant to be a general listing of inputs and parameters that affect model performance as indicated in the first sentence? If the former, reword the introduction. If the latter, add additional inputs and parameters such as the settling velocity of the two flocculated clay/silt classes, and grain size distribution of East River loadings which also affect the performance of the sediment transport model.

  b. Page 125, Bullet List in Paragraph Continued from Page 124: List all the individual model inputs and parameters subject to calibration. Specifically, the settling velocity of two Class 1 classes were developed by calibration, and the composition of three sediment classes from the East River was developed by calibration, making for a total of 5 parameters and inputs that were developed by calibration.

42. Section 7.3 Conceptual Site Models for Hydrodynamics and Sediment Transport

  a. Section 7.3.1, Page 134: The data and model results presented in section 4 are not referenced in the CSM for hydrodynamics described in Section 7. It is not clear what elements of the CSM were developed based on empirical data and what elements were developed using the numerical model. The CSM presented is fairly generic in that it can apply to any small tidal channel. Add text describing the behavior of the system during point source discharge events to describe how these events modify the currents and salinity in the system, and how this could drive the transport of sediment.

  b. Section 7.3.2, Page 135: It is not clear what elements of the CSM were developed based on empirical data and what elements were developed using the numerical model. Consider implementing more empirical lines of evidence in developing the CSM, especially the statements about the relative contributions of East River and point source loadings to sedimentation within various reaches, importance of propwash relative to normal hydrodynamics, potential difference in sediment dynamics during point source discharge events and during dry-weather conditions, etc. Currently, these seem to be based only on model results, but it would be a stronger statement if such findings can be based on empirical measurements. Revise the text accordingly.

  c. Section 7.3.2, Page 135, Third Paragraph: This is the very first mention anywhere in the text on the atypical vertical gradients in TSS during wet-weather versus dry-weather periods. Elaborate on such patterns in Section 5, and add a new sub-section dealing with suspended sediment transport patterns determined from various data-based lines-of-evidence.

  d. Section 7.3.2, Page 136, First Paragraph: See comments to Attachment G-I on the issue of temporal changes in CSO and point source loadings.

Attachment G-F Specific Comments

1. Attachment G-F, Section 1.1 Correlation Analysis of Turbidity and TSS Concentration Data

    a. Page 1, Second Paragraph: Qualify the statement "*Paired samples of turbidity and TSS concentration data were collected during Phase 2 at bulkhead sondes…*". Review of the water depths recorded by the surface and bottom YSI meters show occasional differences in excess of 6' between the depths in the water column where turbidity was measured and where a corresponding water sample was collected for TSS measurements. This is a relatively large difference (relative to the total water column depth), and it also implies that the turbidity and TSS measurements are not truly paired. In other words, measurements can be termed as paired only if made at same point in time and space.

    b. Page 1, Third Paragraph: EPA has reviewed the data used to develop the bulkhead sonde turbidity-TSS correlations shown in Figures G-F-1, G-F-3, G-F-5, G-F-7, G-F-9, and G-F-11. The analysis focused on identifying the sources of variability in the turbidity-TSS relationships. The two major sources of variability include fouling of the turbidity sensors, and differences in the depth sampled by the turbidity sensor and the TSS water sample collection depth. In addition, a smaller subset of water samples also likely include location artifacts, where the water samples were collected in locations with total water depths somewhat different than at the sonde locations.

        Excluding the data affected by the afore-mentioned sources of variability, the turbidity-TSS relationships primarily appear to be a function of the environmental conditions, as seen in Figure 4. The dry-weather relationship includes data from the August 2014 and October 2014 sampling events, and the large (>~80 MG/event) wet weather relationship includes data from the December 2014 and August 2015 sampling events. Data from the remaining events (March 2015, April 2015, and September 2015) consist of relatively smaller wet weather events (~15-35 MG/event). The total point source discharges during these events were estimated using the point source flow input files provided with the FMRM model. The coefficient of determination (R^2) for the dry-weather and large wet-weather relationships are 0.50 and 0.55, respectively. This may partly be due to the fact that variability due to differences in the TSS sampling depth and sensor depths was only reduced (by excluding TSS samples collected at depths more than 3' apart, in the vertical, from the turbidity sensor) but not eliminated entirely. Nonetheless, the individual TSS values are within a +/-2X envelope around the turbidity-TSS regressions, which is typical for such relationships.

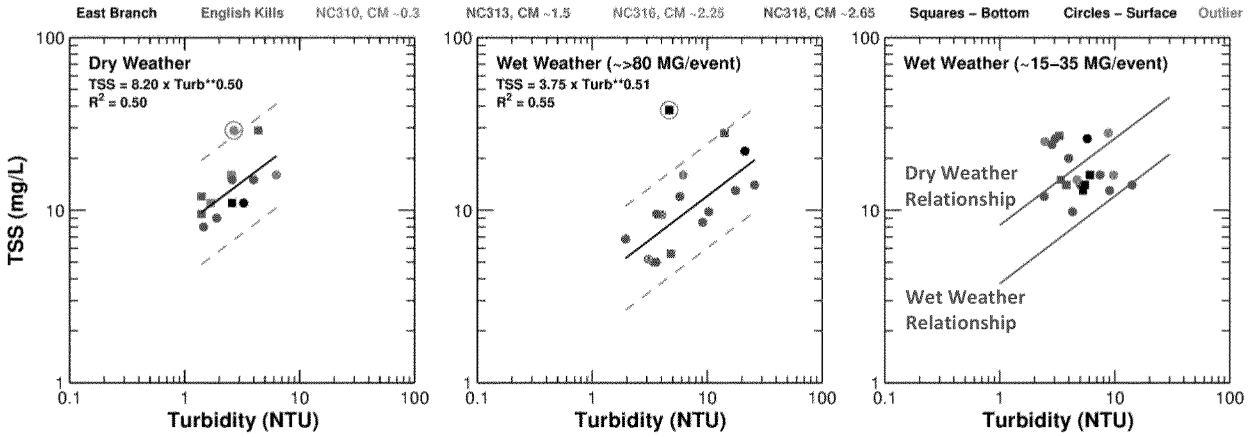

Figure 4. Turbidity-TSS relationships for the bulkhead sondes in Phase 2. Black lines in left and middle panels indicate the turbidity-TSS regression, dashed lines indicate +/- factor of two around the regression. Red lines on right panel indicates the dry-weather and large wet-weather relationships shown in the left and middle panels.

Comparison of the dry-weather and wet-weather relationships shows an apparent increase in turbidity in the entire study area during large wet-weather events. This appears consistent with various aerial images of the study area, which show somewhat more turbid water originating from the tributaries, especially during wet-weather events. An increase in turbidity would also be conceptually consistent with (1) point source discharge events during such conditions, with a higher dissolved organic matter loadings expected from CSO releases than water originating from the East River, and (2) with additional solids loadings from point source releases. Due to the relatively large variability in the turbidity-TSS pairs, a turbidity-TSS relationship was not derived for the smaller wet-weather events (~15-35 MG/event). Nonetheless, it is worth noting that the majority of the turbidity-TSS pairs fall in between the dry-weather and large wet-weather relationships. This is conceptually consistent with the hypothesis of relatively more turbid water associated with point source discharges; releases during the smaller point source discharge events would be subject to relatively more dilution with East River water than during larger events and thus show lesser impacts on turbidity than during large wet weather events.

The dry-weather and large wet-weather turbidity-TSS relationships provide a basis to estimate TSS time-series using turbidity measurements during such conditions. The resulting TSS time-series can provide a basis for understanding sediment transport processes in the system and provide calibration metrics for the sediment transport model under these conditions which also bracket the range of environmental conditions expected in the study area. Include TSS time-series estimated from turbidity data in understanding sediment transport in the system during dry-weather and large wet-weather conditions, and use these TSS estimates in as a model calibration metric.

c. Page 1, Third Paragraph: Refine the regression analyses of the hand-held sondes and measured TSS after reviewing the turbidity-TSS pairs for the various sources of variability noted in EPA's analysis of turbidity-TSS described in the preceding comment:

   i. Turbidity sensor fouling

ii. Consistent sampling depths for turbidity and TSS water samples
iii. Location artifacts, where the water samples may have been taken in a portion of the channel cross-section with significantly different total water depth than the location of the turbidity measurement

2. Section 1.2 Evaluation of NYCDEP and Phase 2 TSS Concentration Data

   a. Page 2, First Complete Paragraph: The comparisons in Figures G-F-14 and G-F-15 make the distinction between near-surface, mid-depth, and near-bottom samples collected during Phase 2. However, no such distinction is made for the NYCDEP stations. Could the NYCDEP data include only near-surface samples or depth-integrated samples? If so, that could explain the difference between the two datasets. Review the data and address in the text and subsequent analyses as appropriate.

3. Section 1.3 ADV and Near-Bottom Turbidimeter Data Collection and Analysis

   a. Page 5, Third Paragraph: Given the apparent increase in turbidity in the system described previously in the discussion of the turbidity-TSS relationships, filtering the turbidity data to exclude periods with large wet-weather discharge events may potentially improve the turbidity-ABS correlations shown in Figures G-F-43 through G-F-48. Filter the turbidity and ABS pairs shown in Figures G-F-43 through G-F-48 by excluding periods of large wet-weather events and reassess the turbidity-ABS correlations, use to estimate TSS time-series, and use for calibrating the propwash scour model as originally described in Section 7.2.3.5 of the Phase 2 RI Work Plan (Anchor QEA, 2014).

   b. Page 5, Third Paragraph: The turbidity-ABS correlation for NC311 shown in Figure G-F-43 has a coefficient of determination (R^2) of 0.66 (this could potentially improve following the suggestion in the preceding comment). This is a potentially useful relationship, at a location experiencing the most navigation impacts of all the ADV deployment stations. Uncertainty in the turbidity-TSS relationship can be incorporated into the TSS estimates resulting from the high-frequency (1 second interval) estimates of turbidity. Review the correlation following the suggestion in the preceding comment, and use for calibrating the propwash scour model as originally described in Section 7.2.3.5 of the Phase 2 RI Work Plan (Anchor QEA, 2014).

   c. Page 5, Third Paragraph: Since the 15-minute turbidity data do not always show the same propwash impacts as the 1-second ADV/ABS data, focusing only on the events where both sets of measurements indicate resuspension may be a defensible approach for evaluating propwash impacts on suspended sediment concentrations. The estimated 1-second interval turbidity can be used to estimate TSS time-series using the dry-weather turbidity-TSS relationship as originally intended in the Section 7.2.3.4 of Volume 2 of the Phase 2 RI Work Plan. The resulting TSS time-series will provide data to calibrate/validate the propwash model described in Section 5.5.6 of Appendix G. Review and develop a strategy to reconcile the 15-minute and 1-second ABS data, and use to estimate TSS time-series for use in calibrating the propwash scour model.

Attachment G-G General Comments

The evaluation of NSRs based on Cs-137 and Pb-210 activity in Section 1.3 proposes historical changes in the sediment loadings from point sources as an explanation for the higher NSRs based on Cs-137 than Pb-210. However, no other lines of evidence (e.g., historical measurements or estimates of point source flows & suspended sediment concentrations, changes in the watershed, point source controls, etc.) are provided in support of the argument of a temporal change in point source sediment loads.

From a conceptual standpoint, there are two constraints on the process of sedimentation – sediment supply and trapping efficiency, with the resulting sedimentation rate a positive function of both constraints. The arguments in Section 1.3 focus only on a hypothesized change in historical point source sediment loadings as an explanation for the temporal decline in sedimentation rate noted in some of the geochronology cores. This process is fairly straightforward – for a given trapping efficiency, sedimentation rate over a given area will be direct function of the sediment loading rate. The limitation of trapping efficiency on sedimentation rate over time can be conceptualized using the approximate geomorphic feedback method used in the FMRM model (described in Section 5.3.4 of Appendix G). For a given flow rate passing a given location in the system, as water depth decreases with sedimentation, velocity increases due to the reduction in cross-sectional area, thus increasing bed shear stress. The increase in bed shear stress causes a decrease in the probability of deposition (calculated using Eq. G-J-7), a parameter linearly related to the sedimentation rate. Therefore, as water depth decreases with increasing sedimentation at a given location, the probability of deposition decreases, thus reducing trapping efficiency, and therefore sedimentation rate. This process is shown graphically, for an arbitrary cross-section with a constant flow rate of 2 m^3/s, constant cross-section width of 10 m, initial depth of 5 m, D90 of 1400 um, and critical shear stress for deposition of 1 dyne/cm^2. With increasing sedimentation (manifest as bathymetric change), flow velocity increases as shown in the upper panel due to a decrease in water depth. This causes an increase in the skin friction at the bed-water interface (also shown in the upper panel). Using the Krone formulation for probability of deposition (one of the commonly used formulations), as shown in the lower panel, the probability of deposition decreases with increasing skin friction (i.e., increasing sedimentation or decreasing water depth). At a bathymetric change of ~4.1 m (corresponding to water depth of 0.9 m), the skin friction becomes equal to the critical shear stress for deposition at which point the probability of deposition reduces to 0 and sedimentation ceases. In other words, the probability of deposition (which is a surrogate for the trapping efficiency) is in a state of dynamic equilibrium with the ongoing sedimentation. This imposes a natural upper limit on the sedimentation that can be achieved in a tidal system such as Newtown Creek. Therefore, the potential for changes in sedimentation rate due to a temporal change in sediment loadings as well as changes in trapping efficiency are to be considered when evaluating data that exhibit temporal changes in sedimentation rate.

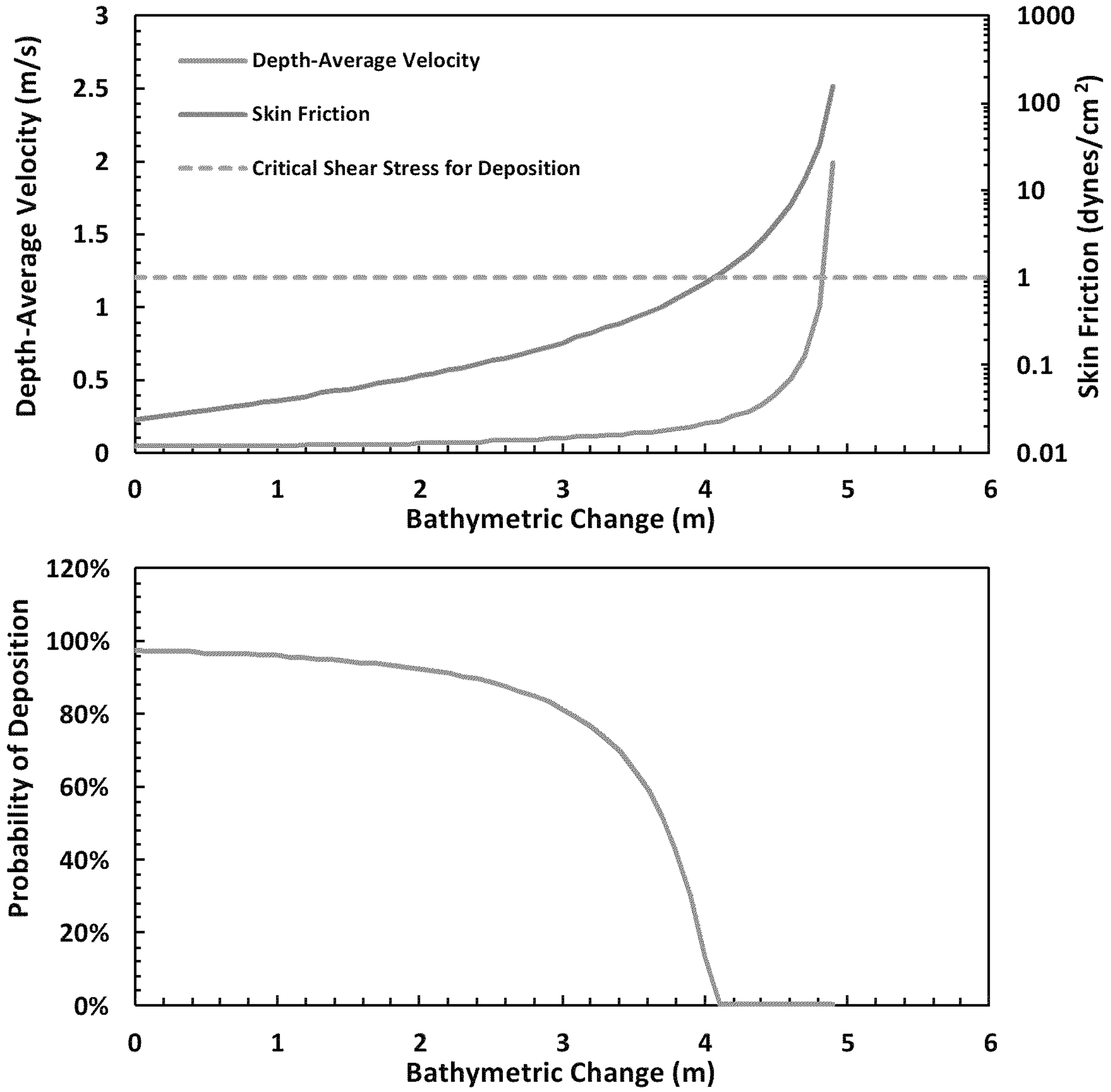

Figure 5. Conceptual depiction of a decrease in the probability of deposition, a surrogate for the trapping efficiency, as a function of increasing sedimentation or decreasing water depth (lower panel). Upper panel shows the increase in depth-average velocity and skin friction as a function of increasing sedimentation.

The geochronology core dataset includes eighteen cores where both cesium-based and lead-based NSRs were calculated. Figure 6 shows the cesium-based versus lead-based NSRs for these cores (left panel), and the cesium-based NSRs versus the 2012 bathymetry at the core locations (right panel). Eight cores (seven in the main stem and one in English Kills) have lead-based NSRs that are similar (within a factor of two) to the cesium-based NSRs. This suggests no temporal changes in sedimentation occurred at these locations, and therefore no changes in sediment supply or trapping efficiency. Review of the remaining ten cores (with cesium-based NSRs more than a factor of two higher than lead-based NSR) relative to the 2012 NCG bathymetry at the core locations suggests that sedimentation rate in a majority of these cores may currently be limited by trapping efficiency. For six of these cores (mostly located within the tributaries), the current bathymetry is relatively shallow and within ~2 ft of Mean Low Water (MLW) levels. Using the cesium-based NSR indicates that in the 1960s, these locations would have been 3 ft to 11 ft deeper than currently. Therefore, a decrease in trapping efficiency may be an equally plausible explanation for the decreasing temporal trend of NSRs in these cores as changes in sediment loadings from the point sources.

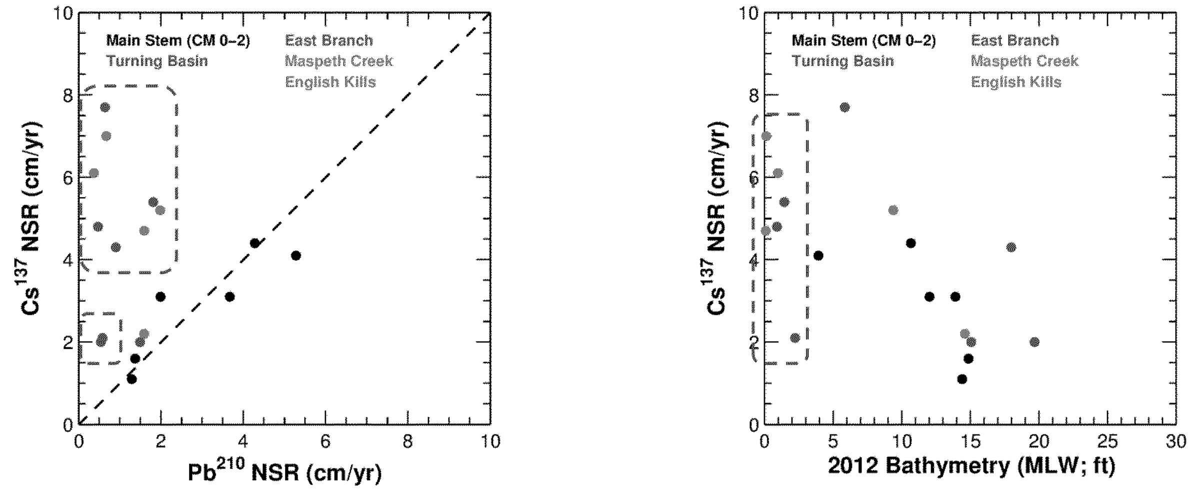

Figure 6. Geochronology NSRs relative to 2012 bathymetry at geochronology core locations.

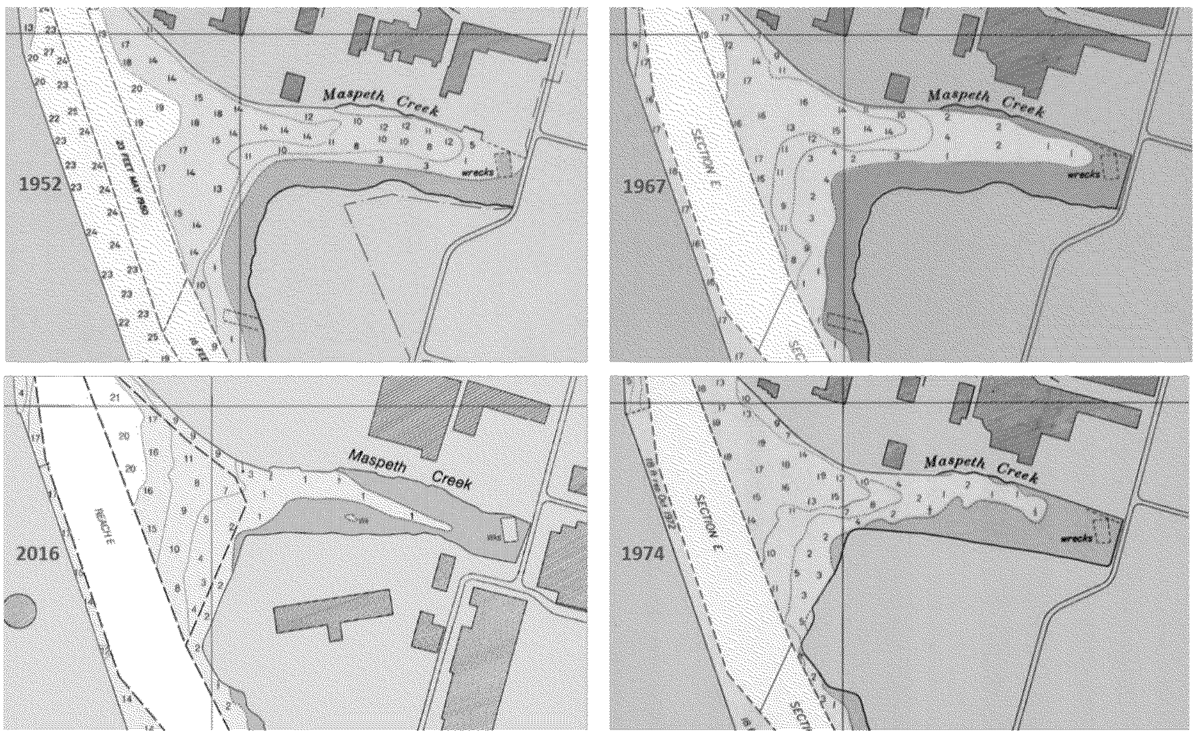

Figure 7. Morphological evolution of Maspeth Creek as seen in NOAA navigation charts. Depths in ft relative to MLW.

The remaining 4 cores (with cesium-based NSRs more than a factor of two higher than lead-based NSR), do not appear to be limited by trapping efficiency based on water depth. However, there may be additional historical changes that may have caused a change in shear stress regime and therefore the trapping efficiency and resulting sedimentation rate. For instance, review of NOAA's navigation charts (shown in Figure 7) along with the dredging history (Section 1.4 in Attachment G-H) of Maspeth Creek shows that even though a navigation channel was dredged in this tributary in the

1930s (to 20 ft below MLW), as seen in the upper left panel in Figure 7, it had infilled significantly by the 1950s. Between 1952 and 1974, the sedimentation rate is on the order of ~10 cm/year. This rapid sedimentation may reflect a change in the shear stress regime due to a decrease in navigation activities (propeller wash can cause a local increase in the shear stress regime thus reducing trapping efficiency) in this tributary sometime between the 1930s and 1952. Since 1974, due to the relatively shallow depths in much of the tributary, sedimentation rate inferred from these navigation charts is on the order of ~1 cm/yr, likely limited by trapping efficiency, due to the relatively shallow water depths. It is also worth noting that the sedimentation rates inferred from these navigation charts is within the range of NSRs calculated from geochronology (both Cs-137 and Pb-210) in this tributary, with the cesium-based NSR approximately ten times higher than the lead-based NSR.

These observations suggest that the difference between the cesium-based and lead-based NSRs cannot be taken solely as an indication of temporal changes in point source sediment loads. There are additional considerations (change in trapping efficiency due to changing water depth as well as navigation impacts) that explain the temporal trends noted in the geochronology NSRs. Revise the text and analysis to include a balanced analysis including alternative possibilities such as temporal changes in trapping efficiency and navigation history that could also partly or wholly explain the observed temporal changes in sedimentation rate.

Attachment G-H General Comments

1. Several analyses related to NSRs using bathymetric data are presented in this attachment. However, in some cases, the findings are not explored in detail. For instance, although the 1991-2012 bathymetric comparison shows mostly sedimentation in the tributaries, the 1999-2012 bathymetric comparison shows erosion over approximately half the area in English Kills and in part of the East Branch (Area 5 in Figure G-H-12). However, these patterns of erosion have not been evaluated from a data quality perspective, i.e. is the pattern of erosion real and explainable given the known forcings in the system (e.g., hydrodynamics, navigation impacts, etc.), or is the pattern of erosion an artifact of a bias in the data.

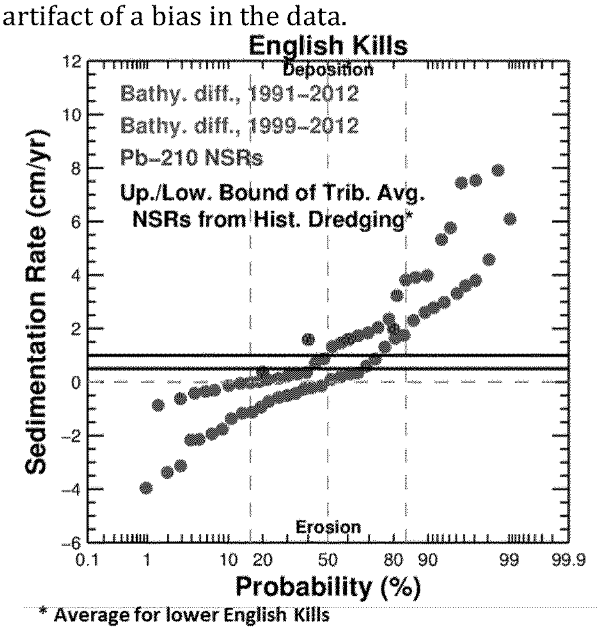

Figure 8 Probabilistic comparison of net sedimentation rate from various lines of evidence.

Figure 8 shows a probabilistic comparison of NSRs in English Kills and represents the results of an analysis to reconcile NSRs from various lines of evidence, assess data quality, and evaluate the erosional signal noted in English Kills during 1999-2012. The NSRs from bathymetric comparison were developed by calculating the difference between the 1991/1999 single-beam bathymetry and the 2012 multi-beam bathymetry at the corresponding locations, with each of the red and blue symbols in Figure 8 representing the average NSR calculated for given single-beam transect. NSRs based on the Pb-210 cores collected in the English Kills are included with green symbols. The upper- and lower-bound estimates of NSR based on historical dredging records (shown in Figure G-H-56) is also included with horizontal black lines.

Figure 8 shows that more than ~50% of the transects show erosion during 1999-2012, whereas only approximately 15% of the transects show erosion during 1991-2012. Spatially, the erosional signal during 1999-2012 is distributed within specific areas within English Kills – Figure G-H-10 shows a consistent pattern of erosion in the southern half of the cross-section in lower English Kills (Areas 3 and 4 shown in Figure G-H-6). In some locations, erosion of up to 6' is calculated over the 1999-2012 period. Comparing the 1991-1999 surveys implies significant sedimentation in these same areas, up to 9' in some cases. The spatially continuous pattern of erosion suggests a distinct signal that is either real or the result of some artifact in the data. However, even if said accumulation from 1991-1999 was explainable (for argument's sake), the erosion signal cannot be explained given the known hydrodynamic and anthropogenic forcings (primarily navigation which is relatively negligible in this tributary at approximately 1 vessel/month). Nor can the erosion signal be attributed to analytical/sampling variability in the 1999 single-beam bathymetry data – such variability would be expected to manifest itself as randomly distributed in space rather than being spatially coherent. Furthermore, all 4 cores with Pb-210 based NSRs were collected in areas where erosion is noted in the 1999-2012 bathymetric comparison. The signal of erosion in the 1999-2012 bathymetric comparison and sedimentation in the Pb-210 dating (which reflects the sedimentation rate over the preceding 10- to 20-year time-horizon, approximately the same time-frame as the 1999-2012 bathymetric comparison) is mutually inconsistent. Furthermore, reviewing the NSRs based on 1991-2012 bathymetric differencing relative to the range of Pb-210 NSRs shows ~50% overlap between the two. In contrast, the Pb-210 NSRs overlap with only roughly 15% of the 1999-2012 bathymetric differencing distribution. In addition, the tributary-average NSR based on historical dredging records is very close to the median NSR based on 1991-2012 bathymetric differencing. These comparisons of NSRs based on 1991-2012 bathymetric differencing, Pb-210, and historical dredging suggests consistency between these various lines of evidence in English Kills and provides confidence in these NSRs. In contrast, the 1999-2012 bathymetric differencing appears to be an outlier, showing erosion that cannot be explained given known forcings, and is inconsistent with the Pb-210 and historical dredging NSRs. Therefore, the 1999-2012 bathymetric comparison may likely be unreliable, possibly due to artifacts in the 1999 bathymetric survey.

Similar analysis comparing the 1991-2012 bathymetric differencing, 1999-2012 bathymetric differencing, Pb-210 NSRs, and NSRs based on historical dredging records for the other tributaries (East Branch, Maspeth Creek, and Dutch Kills) show, for the most part, consistency between the various lines of evidence. Exceptions include:

a. The 1999-2012 bathymetric differencing in East Branch which seems to be affected by the same artifact as the English Kills, although to a smaller degree
b. Uncertainty originating from Pb-210 NSR in one of the cores in Dutch Kills
c. A localized difference in sedimentation rate near the mouth of Maspeth Creek, higher in the 1991-1999 time-frame than in the 1999-2012 time-frame.

The review of the NSRs from various lines of evidence suggests a data quality issue in the 1999 bathymetry leading to unexplainable results in the 1999-2012 bathymetric differencing in English Kills and to a minor extent in the East Branch. This suggests that the 1999 bathymetry should not be used to support the modeling efforts in these tributaries. In contrast, within Maspeth Creek, the 1999-2012 bathymetric differencing is consistent with the 1991-2012 differencing over the majority of the areal extent of the tributary, the only exception being an area of high infill located near the mouth of the tributary. For the most part, the NSRs based on 1991-2012 in English Kills and East Branch, 1991-2012 and 1999-2012 in Maspeth Creek, and 1999-2012 in Dutch Kills are consistent with NSRs based on Pb-210 and historical dredging records.
Perform a comparative analysis of NSRs from the various lines of evidence as a data quality check and to reconcile NSRs from various lines of evidence.

<u>Attachment G-H Specific Comments</u>

1. Section 1 Estimation of Net Sedimentation Rates Based on Differential Bathymetry Analysis:

   a. Page 1, First and last bullets: Given the availability of 1999 bathymetry in Dutch Kills, either (1) include Dutch Kills in this analysis, or (2) provide justification as to why Dutch Kills is being excluded.

2. Section 1.1 Differential Bathymetry Analysis: 1999 to 2012

   a. Page 2, Second paragraph: Clarify the description in this paragraph. It suggests that in the near-shore zone, the 2012 multi-beam bathymetry data consists of single-beam data from 2011 and LiDAR data, then discusses uncertainty in the 2012 data in the near-shore zone. While a combined dataset may have been generated and referred to as the 2012 bathymetry, in the interest of clarity, suggest developing some alternative terminology to refer to this combined dataset. The 2012 multi-beam bathymetry should refer to only the multi-beam bathymetry collected in 2012. Any combination of this dataset with data from other years should be termed appropriately in the text.

   b. Page 3, Third paragraph: Based on the values in Table G-H-1 and areal extents in Figure G-H-6, approximately half the length of English Kills experiences net erosion during 1999-2012. However, neither the text in Appendix G nor the text in Attachment G-H discusses this pattern of erosion. If real, it represents the only truly observed erosion signal in the system, and has important implications for the fate and transport of sediments and contaminants, with the eroded sediment (and contaminants) potentially depositing elsewhere in the system. Furthermore, this erosion occurs over the duration of the sediment transport model calibration period (1999-2012) and is therefore a feature that should be reproduced by the model. Revise the text to discuss in detail the various features noted in the bathymetric difference data along with likely mechanisms that may explain the noted patterns of erosion and deposition.

3. Section 1.2 Differential Bathymetry Analysis: 1991 to 2012

   a. Page 4, Second paragraph: In addition to Dutch Kills and portions of Whale Creek, 1991 data is unavailable also in portions of the East Branch. Revise the text accordingly.

   b. Table G-H-2: The area-average NSR for Maspeth Creek seems wrong. Comparing to Table G-H-3 suggests that value in Table G-H-2 is only for Area 1 in Maspeth Creek rather than the entire tributary. Revise the table as appropriate.

   c. Page 4, Second paragraph: The second sentence of this paragraph suggests that Table G-H-3 includes a comparison of 1991-2012 and 1999-2012 NSRs by sub-area for the tributaries. However, Table G-H-3 only includes a tabulation of the 1991-2012 NSRs by sub-area for the tributaries. Correct the text to accurately reflect the contents of Table G-H-3 or update the table to be consistent with the text.

   d. Table G-H-3: NSR for Area 4 over 1991-2012 should not be included to the extremely limited bathymetry coverage in 1991 (see left panel on Figure G-H-45 for 1991 coverage). Exclude Area 4 for 1991-2012 from Table G-H-3.

   e. Figure G-H-46: With the exception of Area 2, the remaining areas in English Kills exhibit either net erosion (areas 1 and 4) or very limited sedimentation (area 3) over 1999-2012. The temporal changes in behavior inferred during 1991-1999 (net accumulation) and during 1999-2012 (net erosion) in areas 1, 3, and 4 are not discussed in the text. Given the fact that these represent the only observed signal of erosion within the study area, revise the text to discuss them in further detail along with potential mechanisms that may explain the measured erosion signal.

   f. Figure G-H-47: NSR for Area 4 over 1991-2012 should not be included due to the limited bathymetry coverage in 1991 (see left panel on Figure G-H-45 for 1991 coverage). Exclude Area 4 for 1991-2012 from Figure G-H-47.

4. Section 1.4 Differential Bathymetry Analysis: Historical Dredging Periods to 2012

   a. Page 6, First paragraph: Provide references for the historical dredging data.

   b. Figure G-H-50: Describe how the year of last dredging was developed, in particular, the spatial distribution. For instance, within the main stem, a ~0.05-mile section around CM 0.4 is shown as being dredged in the 1930s even as areas immediately upstream and downstream are shown as being dredged in the 1940s. Similar areas are also seen along the southern shoreline in the Turning Basin, the entrance to Maspeth Creek, and just upstream of the Turning Basin. This figure also shows a ~0.05-mile stretch between the Turning Basin and English Kills/East Branch where no dredging is shown to have ever occurred, which seems unusual given that areas upstream and downstream of this stretch were dredged.

c. Figure G-H-51: A 16' dredge depth shown for a ~0.05-mile stretch between the Turning Basin and English Kills/East Branch is unlikely considering that areas upstream were dredged to 18'. Either revise Figure G-H-51 with an 18' dredge depth for this area or provide evidence for a 16' dredge depth.

d. Page 6, Second paragraph, third sentence: The fact that current depth is greater than target dredging depth does not necessarily mean erosion occurred since dredging. It could also mean those areas were naturally deep and were therefore not dredged during the last dredging event. Note this uncertainty in the inferred pattern of erosion/deposition in the text and in the resulting NSRs shown in Figure G-H-56.

5. Section 1.5 Differential Bathymetry Analysis: 1999 to 2011

   a. Figure G-H-67: X-axis labels are missing. Revise the figure.

Attachment G-I General Comments

6. This attachment presents the results of a mass balance analysis focusing on potential temporal changes in sediment loadings from point source discharges. The focus of the analysis is on quantifying the point source sediment loadings over the 1991-1999 and 1999-2012 time-frames. However, the purpose of this analysis does not seem to be directly related to the model as applied in the draft RI. The model is applied over the period 1999-2015, and therefore the issue of temporal changes in point source loadings before and after 1999 is inconsequential to model development and calibration. Furthermore, the application of the model during the feasibility study will be to future conditions, and therefore historical point source loadings are of no interest from that perspective either. Therefore, it is not clear why an analysis to quantify potential temporal changes in point source loadings is necessary in the RI report. Either make the connection to the RI in the text or remove this attachment entirely.

Attachment G-I Specific Comments

1. Section 1 Data-based Sediment Mass Balance Analyses:

   a. Page 1, First Paragraph: As described in the text, the analyses was performed only for English Kills, East Branch, and Maspeth Creek. However, in the Dutch Kills, despite the availability of bathymetry data from 1999 and 2012, as well as Pb-210 and Cs-137 based NSRs, the text makes no mention of this tributary. Include either (1) a justification of why this analysis was not performed for Dutch Kills, or (2) the results of such an analysis for the Dutch Kills.

   b. Page 1, Second Paragraph: Add justification for the statement "*Average point source sediment loads during the 14-year calibration period were likely higher than sediment loads during the 2015 point source sampling period, due to decreasing combined sewer overflow sediment loads during the calibration period.*"

   c. Page 1-2, Paragraph 5 starting on page 1, and First and Second Paragraphs on Page 2: The large difference in NSRs between the 1991-2012 and 1999-2012 bathymetric comparisons may be an artifact of a potential data quality issue affecting the 1999 bathymetry survey, primarily in the English Kills and to a minor extent in the East Branch as well. See comments to Attachment G-H regarding potential data quality issues affecting the 1999 bathymetry. Review and revise the text as appropriate.

d. Page 2, Last Paragraph: In order to accept the hypothesis that point source sediment loads in the 1991-1999 time-period were greater than during the 1999-2012 period, provide additional supporting evidence. Such evidence can be in the form of measured flows and/or suspended sediment concentrations from the point sources, implementation of watershed-level best management practices, changes to the operation of the sewer system, wastewater treatment capacities, etc., that may have altered the point source solids load over time. As such, the analysis of the 1991, 1999, and 2012 bathymetric data can only be treated as indirect evidence. Other possibilities may also explain a higher sedimentation rate in the 1991-1999 time-frame as compared to the 1999-2012 time-frame. For instance, it is possible that relatively large navigation impacts within the main stem during the 1991-1999 (for instance more traffic during 1991-1999 than during 1999-2012) time-period may have limited accumulation in the main stem and caused relatively large (compared to conditions after 1999) net up-creek transport of sediments to the tributaries. Direct evidence in support of temporal changes in the point source loadings will support and strengthen what is at this point one hypothesis that could explain the temporal change in sedimentation rate.

e. Page 3: The statement "*the minimum point source sediment loads in these three tributaries for the 1991 to 1999 period correspond to the data-based mass deposition rates for that period*" is not definitive since it ignores the possibility of sediment originating from downstream locations and transported into the tributaries and depositing. Revise by either (1) listing the assumption of no solids from downstream depositing in the tributaries, or (2) including adequate justification of why no solids from downstream would have deposited in the tributaries.

f. Figure G-I-13: Include values corresponding to the 1999-2012 and the 1991-2012 periods on both panels.

Attachment G-J Specific Comment

1. The model application for Newtown Creek uses the Partheniades formulation for probability of deposition. Include this formulation a part of the model formulations documented in this Attachment.

Attachment G-L General Comment

1. Although the analysis in Section 1.1 in this attachment are referred to in the text of Appendix G, Section 1.2 in this attachment is not referenced in the text of Appendix G. Either (1) provide such reference in the main body of the text, or (2) delete this Section 1.2 from this Attachment.

Attachment G-M General Comments

1. This attachment is currently included in Appendix G without any reference to the text in Appendix G. Either (1) provide such reference in the main body of the text, or (2) delete this Attachment

2. Section 1.3 Effects of Bed Consolidation on Predicted Net Sedimentation Rates, Pg 3, 2nd paragraph. Revise the report to explain in detail any adjustment that was made to NSRs for deeper sediments.

References

Anchor QEA, 2014. "Phase 2 Remedial Investigation Work Plan – Volume 2. Remedial Investigation/Feasibility Study, Newtown Creek", Woodcliff Lake, NJ.

Fugate, D and R Chant, 2006. "Aggregate Settling Velocity of Combined Sewage Overflow", Marine Pollution Bulletin, 52, pp 427-432.

Hayter, E.J. 2016. Evaluation of Anchor QEA's Preliminary Modeling Results Memorandum and the Phase 1 Hydrodynamic and Sediment Transport Models. Letter Report. U.S. Army Engineer Research and Development Center, Vicksburg, MS.

HydroQual, 2007. "A Model for the Evaluation and Management of Contaminants of Concern in Water, Sediment, and Biota in the NY/NJ Harbor Estuary: Sediment Transport/Organic Carbon Production Sub-Model", Contamination Assessment and Reduction Project (CARP), Prepared for the Hudson River Foundation, NY.

Louis Berger Group (LBG), HDR, and Battelle, 2014. "Focused Feasibility Study Report, Lower Eight Miles of the Lower Passaic River", Prepared for the US EPA Region 2, and the US Army Corps of Engineers.

Ralston, DK, JC Warner, WR Geyer, and GR Wall, 2013. "Sediment Transport due to Extreme Events: The Hudson River Estuary after Tropical Storms Irene and Lee", Geophysical Research Letters, Vol. 40, 5451–5455, doi:10.1002/2013GL057906.

STOWA/RIZA, 1999, Smooth Modeling in Water Management, Good Modelling Practice Handbook; STOWA report 99-05, Dutch Dept. of Public Works, Institute for Inland Water Management and Waste Water Treatment report 99.036, ISBN 90-5773-056-1

US ACE, 2013. "Hydrographic Surveying", EM 1110-2-1003, November 2013.

US EPA, 2009. "Guidance on the Development, Evaluation, and Application of Environmental Models", Office of the Science Advisor, Council for Regulatory Environmental Modeling, EPA/100/K-09/003, March 2009

US EPA, 2010. "Modeling Journal", New England Regional Laboratory's Office of Environmental Measurement and Evaluation - Quality Assurance Unit, May 2010